The Indo-European Language Family: Questions about its Status

Edited by:

Angela Marcantonio

Journal of Indo-European Studies Monograph Series, No. 55
Institute for the Study of Man
Washington, DC
2009

The Indo-European Language Family:

Questions about its Status

Edited by:

Angela Marcantonio

Journal of Indo-European Studies Monograph Series, No. 55

Library of Congress Number 2009920649
ISBN 978-0-941694-02-5

Contents

INTRODUCTION

Angela Marcantonio
University of Rome 'La Sapienza'

1. 1. The purpose of the present volume is to survey the current state of the Indo-European ('IE') theory in the light of modern linguistic knowledge. Included in the survey is also extra-linguistic evidence, such as recent archaeological, genetic and palaeo-anthropological findings. Its ultimate purpose is to revisit the validity of the various tenets of the theory. In fact, when scholars refer to the "IE theory", they may be referring to one of a number of competing, and sometimes contradictory, models. For example, some regard IE as purely a linguistic classification, whilst others regard it as an attempt to reconstruct a 'real' pre-historical language. To take another example, some scholars hold that the original IE protolanguage was a morphologically complex language, similar to Sanskrit, whilst others argue it was morphologically simple.

The volume intends to achieve its purpose by approaching the IE theory from various perspectives and areas of expertise, thanks to a collection of articles by scholars specialized in IE and historical linguistics, and also by Sanskritists, Dravidianists and archaeologists, whose field of study have been heavily affected by the IE linguistic classification. It is only by bringing together, in a single volume, these various approaches and views about IE that it clearly and explicitly emerges how surprisingly different and, often, even deeply contradictory, these views and approaches may be. Thus, the IE theory is widely accepted despite the fact that opinions differ enormously on what the theory actually comprises. Opinions may clash even as to the very nature and validity of many of the underlying tenets of the theory – whether explicitly stated or quietly assumed. For example, as mentioned, some scholars regard the subject a 'pure theory', which helps to describe correlations between languages, whilst others regard it as a valid means to reconstruct pre-historical facts. Although there seems to

be a widely shared assumption that at least a 'hard core' of the correlations among the IE languages are 'compelling' – that is, too striking to be the result of chance – there may be deep divergences on how to interpret these correlations, as well as on other, less central but equally important aspects of the theory. In other words, contrary to what one would expect, the wide acceptance of the IE linguistic classification does not appear to be accompanied by a parallel acceptance of a coherent and equally agreed set of tenets, a coherent common denominator, consisting of what one could call the 'hard linguistic evidence' and the 'fundamental principles' upon which the theory, supposedly, is based. This being the case, it is appropriate to dedicate a volume to the precise task of addressing the fundamental question of why this should be the case. This, ultimately, amounts to the task of re-assessing the founding principles of the theory. In particular, one might reasonably ask the following questions:

1. How is it possible that such widely differing and often contradictory views are still unresolved today? Is this a consequence of the fact that the adopted methods of analysis are not rigorous enough, or alternatively is it the linguistic and extra-linguistic evidence that is problematic – for example, is it ambiguous, insufficient, or simply highly 'malleable'?
2. Is the IE theory a 'scientific' theory, that is, has it been established through "methods of analysis which produce results that are subject to invalidation"[1]? If so, what precisely is the evidence counter to the model? This evidence does exist, although it is usually minimized, in various ways, but: is it really so insignificant?
3. As an alternative, some scholars believe that the theory cannot be invalidated and therefore is not a 'scientific' theory; nevertheless its validity is simply 'compelling'. What is the basis for this claim?

The points of views and perspectives presented in this volume, although might have surfaced in professional publications before,

have hardly been dealt with and confronted with one another within a single volume, despite being inextricably interdependent. As a consequence, the questions raised in the points (1)-(3) above have hardly been addressed in a targeted and systematic way. For example, there are plenty of publications dealing with the issue of the strengths and weaknesses of the methods of historical linguistics (see Birnbaum (1977); Jones (1993); Fox (1995); Aikhenvald & Dixon (eds, 2001); Nichols (1992); Lass (1997), etc., just to pick up some names at random). However, these publications hardly ever ponder on whether the acknowledged weaknesses may have a negative impact on the IE theory and, if yes, to what extent. Similarly, there are many publications revolving around "how real(ist) are reconstructions" (Lass 1993), but their scope hardly ever extends to encompass the consequent issue of how to best interpret the IE reconstructed, comparative corpus. On the other hand, there are plenty of publications which (ignoring the debates mentioned above), revolve only around the question of the whereabouts of the (assumed) IE proto-community. Finally, textbooks of IE linguistics (and, often, specialist publications too), hardly ever mention any of these ongoing debates. It is true to say that the purpose and scope of a textbook is not that of reporting controversies, but just that of presenting and illustrating the received wisdom. Nevertheless, the fact remains that textbooks typically present a highly idealized, monolithic picture of IE: a paradigmatic, problem-free language family, where everything works (especially sound laws, lexical and morphological correspondences), where there are hardly any contradictions or ambiguities in the linguistic or extra-linguistic evidence, or even any significant divergence of views among scholars. This is even more surprising since this idealized picture is in stark contrast with the messy reality – in terms of variation, high level of exception, contradictory evidence, etc. – found in practically all the other language families of the world, as well, admittedly, within branches of IE (such as the Romance and Germanic languages, or the Balto-Slavic continuum, etc.). This in turn has lead several scholars (such as Grace 1990 & 1996) to come up with the rather 'aberrant' idea that there must exist in the

world two basic types of language families: the 'the exemplary ones' (IE and just few others) and the 'aberrant' ones (all the rest).

There are, of course, textbooks that present a more realistic picture of the state of the IE family, by pointing out the many exceptions to the stated laws, the difficulties encountered in the process of phonological or morphological reconstruction, or the poor quality and /or quantity of the evidence in support of otherwise widely accepted theories. One may compare, for example, Szemerényi's (1973 & 1996:122 ff.) and Gusmani's (1979) account of the slippery evidence on which the laryngeal theory is typically based on the Hittite side, or Sihler's (1995:144 ff.) account of the factual and methodological obscurities found in Verner's Law, one of the most revered IE sound laws. However, in most cases, the dissident opinions or the problematic data and analyses – if mentioned at all – are typically minimized through various means (for example, by being reported in footnotes and / or by the absence of the due references), so that it is difficult to evaluate their potential impact on the validity of the theory, or simply just to acknowledge them. It could be objected that no theory can be or ever has been invalidated by the discovery of one or more pieces of evidence counter to the model. This is true; however, if, at some stage in the history of a theory the amount of evidence counter to the model reaches what is usually called a 'point of critical mass', then a revision of the theory might be in order, whether to modify or update its tenets, or, if necessary, to reject it altogether. In other words, minimizing, re-interpreting or adjusting any evidence found to be inconsistent with the model might be misleading, and therefore not desirable. In fact, scholars identifying problems in their area of research may wrongly assume that the matter has been settled beyond doubt in other areas of study, and may therefore in turn decide (understandably) to minimize or even ignore his / her own item(s) of counter-evidence, in this way, unwillingly, and maybe wrongly, contributing to reinforce the validity of the theory in question.

1. 2. The scholars contributing to this volume will argue in favor or against some of the major tenets embedded in the variegated IE

theory on the basis of both specific corpora of data and / or methodological considerations. We believe that a fair and healthy debate may be of help in attempting to sort out, if necessary, what appear to be proper 'facts' and valid analyses from what appear to be instead (questionable) interpretations, (unfounded) speculations, or even sheer myths. If it turns out from this debate that the different and at times contradictory views about IE can, after all, be reconciled, then the IE theory will gain in rigor and therefore in credibility. If, on the contrary, it turns out that these views cannot be reconciled, then, maybe, some or all of the tenets and assumptions that lie at the foundation of the theory will have to be revisited, or called into question.

This volume sets out to bring together points of view and analyses that deal with issues within the following, main pillars of IE studies: a) the sound laws and the reconstructed phonology and lexicon; b) the reconstructed grammar; c) the traditional family tree model and related concept of 'proto-language' *vs* other models of language formation and diffusion; c) the thesis that the Sanskrit language /culture did not originate in India, but represents an intrusion from the west; d) the thesis that Sanskrit is (arguably) the oldest language within IE, and therefore one of the most, if not the most important one for the purpose of reconstruction. In addition, some attention will also be dedicated to the state of the archaeological, palaeo-anthropological and (to a lesser extent) genetic research, to verify whether or not the 'extra-linguistic evidence' lends support to the traditional model (although this is a linguistic volume and its focus remains the linguistic analysis). As to the specific areas of debate, we will concentrate on the following: 1) the strengths and weaknesses of the comparative method, and, therefore, the reliability of the conventionally established sound laws and reconstructions; 2) the role of morphology in assessing and reconstructing language families in general and the IE family, with particular attention being paid to the IE verbal system; 3) the debate of whether it is possible to tell apart genetic from areal correlations; 4A) the 'conventionalist' *vs* 'realist' approach to reconstruction; 4B) the issues of the IE

homeland and related migrations of the (assumed) IE proto-community.

2. The comparative method and the sound laws

2.1. The strengths and weaknesses of the comparative method

The debate revolving around the issue of the strengths and weaknesses of the comparative method is an intense, long standing debate. In fact, establishing regular sound correspondences is considered by several (many / most?) scholars to be a crucial part of the process of establishing language families (see for example Campbell (1998:315)). However, as mentioned, this debate has hardly ever been associated with a targeted, extensive investigation of the possible impact the weaknesses of the comparative method may have on the validity of IE as a linguistic classification. In particular, the long standing "*Lautgesetz* controversy" (for which see Wilbur 1977) subsided without resolution, and despite its recent resurfacing in publications dealing with several linguistic areas / families in the world (see for example Ross & Durie (eds, 1996); Blust (1996)[2], Aikhenvald & Dixon (eds, 2001), etc.), it is rarely referred to in textbooks of IE. In fact, these usually assume, whether explicitly or implicitly, that the 'regularity principle' and the related family tree model have been established in IE beyond doubt, through the support of an extensive amount of data derived from the various IE languages. However, this is not necessarily the case (as illustrated in this volume in the chapter by Andersen and Marcantonio). As a consequence, the reader, including general linguists or even experts in historical linguistics who are not acquainted with the details of IE, may be excused if they are confused as to the actual 'quality' and 'quantity' of the phonological / lexical correspondences conventionally established for IE. Indeed, within IE too irregularity and exceptions do occur – it would be odd otherwise. However, the vital questions are:

A) How pervasive is this irregularity?

B) Is it really true that the encountered irregularities can, in most cases, be justified through 'genuine' linguistic processes, that is, without stretching the explanatory system up to the point of dangerous 'circularity', by 'appending' *ad-hoc* justifications?

The reader who would expect an answer to these questions which is coherent, unanimous, and, most importantly, decisive – one way or the other – might be disappointed. To show that this is the case, here is a list of the most common justifications put forward to answer question (A):

1. Only the ordinary nouns in IE, particularly those referring to objects and concepts of everyday life, display a high degree of irregularity (much higher in any case than verbal roots), but this is only because they belong to the lower level of speech, the lower level stratum of the IE population (Meillet 1934: 396 ff.; Benveniste 1935:175 ff.).
2. The irregularities are only, in most cases apparent, in the sense that linguists have not yet found the appropriate explanation to account for them.
3. There are some / many / plenty of irregularities (according to interpretations), but there is nothing to be surprised about. We know that sound changes do not proceed so regularly after all, but this does not have a negative impact on the validity of IE, whose establishment, in fact, has *not* been based (only) on the phonological / lexical correlations, but on the morphological ones. Such a view is embraced, among others, by Harrison (2003:214 ff), Greenberg (1987:18; see also Croft 2005) and, in this volume, by Kazanas. For example, Harrison (2003: 217) states that: "If one can *prove* that even one single cognate pair holds over two languages, one has proven those languages genetically related". The (implicit) claim here appears to be that proper lexical 'correspondences' are generally hard, if not impossible to attain. Therefore,

linguists must come to terms with the fact that also within IE (like within other language families or across macro-families) the correlations among the assumed cognates are not, in the main, as 'regular and systematic' as generally claimed.

4. Sound changes do, in the main, proceed regularly, but the encountered irregularities are a natural effect of the great antiquity of the IE family, although its precise degree of antiquity is impossible to assess (see in this regard the chapters by Kazanas, Bryant and Schmitt in this volume).

One may further observe that sound changes (such as phonemic mergers) may not be necessarily that useful for the purpose of identifying innovations, and therefore the internal sub-grouping of families. Most of these sound changes are in fact so natural that they can be easily repeated in different lines of descent[3]. Besides, these natural sound changes may independently occur in totally different language families (see for example Ringe et al. (2002:63 ff.); see also Blenvins (2004) for a detailed account of word-wild frequent sound patterns and sound changes[4]).

As to the answer to question (B), we are not aware of any research carried out with the specific purpose of ascertaining whether or not the circularity issue has had a negative impact on the soundness of IE. In the year 1998 Morpurgo Davies (1998: 254) wrote:

> A final agreement about the nature and validity of sound-laws was never reached. It was generally accepted (by the neogrammarians and everyone else) that testing any sound law against the data was bound to reveal a number of exceptions; in other words, there could not be an immediate empirical demonstration of the regularity principle [...]. The neogrammarians did of course maintain that all the exceptions could be explained away by re-defining the law, or by identifying a different starting-point, or by recognizing the interference with analogical process, but they were immediately accused of circularity [...]. We can say that the sound laws have no exceptions only because when we find an exception we eliminate it saying that there has been analogical interference. On the other hand, we also say that the only way in which we can

> prove that a form [...] is analogical is by pointing out that otherwise it would be an exception to the sound laws. There seemed no obvious counter-objection and the problem remained open [...]. Meanwhile, however, the practicing historical or comparative linguist continued to, or began to, operate in terms of sound laws not too remote from those pleaded by the neo-grammarians.

On the issue of the reliability (or otherwise) of sound laws, compare also the following statement made by Clackson (2007: 60-61) with specific reference to the laryngeal theory: "The comparative method does not rely on absolute regularity, and the PIE laryngeals may provide an example of where reconstruction is possible without the assumption of rigid sound-laws". This statement begs the question of why, when and where, and on the basis of which criteria, scholars may – or may not – assume the existence of "rigid sound-laws".

2. 2. The circularity issue

If – as it seems – the circularity issue has not been solved (yet?), scholars could attempt to set up some sort of qualitative and / or quantitative constraint to the number of the defining parameters a given law may consist of. In practice, however, as far as we are aware, there has never been any such attempt. On the contrary, in the every-day, painstaking practice of establishing sound laws and correspondences, any mismatch in the evidence (ambiguity or absence of the expected outcome, exceptions, etc.) can always "be explained away" through a range of procedures, a range of 'adjustable parameters', to be added to the original definition of the law. In other words, instead of casting doubts on the validity of an assumed law (and, if necessary, dropping it) when faced with exceptions and difficulties, typically the practitioner tries to 'rescue' the stated law, even at a cost of making recourse to a (virtually unlimited) number of (often *ad-hoc*) adjustments, such as:

1. re-defining the law
2. identifying a different starting-point of the law

3. assuming borrowing, from other languages, or 'transitional' dialects, or even from unknown, extinct languages /dialects[5]
4. assuming analogical processes
5. re-arranging the stated sequence of rules in a different order
6. postulating a (/another) laryngeal segment
7. stating that the mismatches in the expected outcomes of the law are not significant for calling into question a theory as well established as IE.

The obvious consequence of this circularity deeply embedded in the comparative method is that the explanatory system runs into the risk of becoming so powerful, so flexible, that it can be stretched to match almost any data, in this way making it impossible to compare the results it yields against the predictions of the model. In other words, although each single 'adjustable parameter' as listed above may in itself reflect a plausible, genuine linguistic process, the overall cumulative effect of many adjustable parameters added to the definition of a given law may endanger the 'cumulative effect', the 'statistical significance' any established 'law', or even 'tendency', should display to deserve these names. This is an issue that has hardly ever been properly and systematically addressed (as far as we know). Although the supposedly rigorous, 'scientific' nature of the comparative method has often been called into question, and more objective quantitative methods of analysis have been at times adopted within historical linguistics, the statistical significance of the IE comparative corpus itself (both the phonological (/lexical) corpus and the morphological one) has never been tested. For example, Collinge (1985) raises the general issue of the lack of 'cumulative effect' in most of the conventionally established 'Laws of Indo-European', but does not provide any detailed, systematic (qualitative and / or quantitative) analysis of the IE comparative corpus on the whole. Ringe (1992, 1995, 1998a, 1999a) and Ringe et al. 2002) apply their statistical analyses basically to the correlations between IE and other, supposedly related language

families, in order to prove that macro-families lack statistical significance. Other authors, such as McMahon & McMahon (2003 & 2005), or McMahon (2005), apply quantitative analyses to estimate the 'correct phylogeny' of the IE family, whilst also emphasizing the need of integrating the traditional methods of historical linguistics with quantitative methods (at this regard see the chapter by Drinka, who also discusses the pros and cons of the latest cladistic models). Thus, these Authors have used the IE family, whose validity is taken for granted, basically as a 'control case' for various kinds of statistical tests within historical linguistics, but have not tested the statistical significance of the IE comparative corpus itself. See in this regard Brady & Marcantonio (2003) and Marcantonio (2002/5). See also Marcantonio in this volume, who argues that the great majority of the conventionally stated IE sound laws lack statistical significance and that, therefore, most of the conventionally established correspondences (within the **LIV** corpus) are simply similarities, most probably in the given sense of 'chance resemblances'.

2. 3. Borrowing *vs* inheritance

In addition to the issues listed above, there is also to consider the possibility that the established 'cognates' – be they 'similarities' or proper 'correspondences' – may be due to the common processes of borrowing, diffusion, convergence, or even chance resemblances. As is known, borrowed words tend to integrate into the sound pattern of the receiving language, as well as undergoing the same (more or less regular) changes that inherited words would undergo. Thus, the identification of borrowed elements on the basis of internal, linguistic clues only might not always be easy. Therefore, sound correspondences, whilst fundamental to most approaches in assessing language families, "can be misused" (to use Campbell's (1998: 315) words). The difficulties (and related long standing debate) of telling apart genetic *vs* areal correlations are clearly illustrated in this volume in the chapter by Andersen, who closely examines certain pathways of development of the Balto-Slavic accentual system.

It is also worth mentioning in this context the well known fact that several semantic fields within the IE basic lexicon (including the terms designating 'extended family' and 'the family of the woman / wife'), in addition to being mainly irregular, typically lack a wide distribution across the IE area, being often confined to just two or three contiguous languages (as pointed out by Schmidt (1872) and Meillet [1934]). In contrast, the cognate terms for 'mother', 'father', 'brother', 'daughter', etc., display a much wider distribution, and a higher degree of regularity. This factor has correctly raised the suspicion that processes of loss (and consequent replacement) of original words, or even processes of chance resemblance, may have been involved in this area. The issue of the wide *vs* restricted distribution and the (degree of) irregularity of many basic cognates within IE is interpreted differently by different scholars (see for example discussion at point (1) in paragraph **2.1.** above). On this topic the reader may compare the contrasting views held by Kazanas and Marcantonio in this volume. Kazanas[6] argues that a great deal of fundamental cognate terms have actually been individuated across the IE languages, more than enough to warrant their inherited nature – even if they are not, in the main, regular and systematic correspondences. Marcantonio, on the other hand, argues that there is a great deal of chance resemblances involved. Finally, one could also assess these data and related interpretation in the light of recent research in the field of historical, contact linguistics, for which see Matras (2002) and Matras & Bakker (eds, 2003). The results of this research appear to support the 'loss' interpretation. In fact, Matras (2002) investigates two typical mixed languages, Romani and Domari (descendent of Central Indo-Aryan languages) and shows how patterns of original words are typically retained, and are therefore easily identifiable even within a context of extensive borrowing. For example, the system of kinship terms of Romani generally retains original, Indo-Aryan terms that refer to "the first-level kin of the same generation" (such as the terms for 'sister' and 'brother'), or the "first-level and lateral kin one generation older" (such as the terms for 'father', 'mother', 'uncle', etc.), or the "first level kin one generation younger" (such as the

terms for 'son' and 'daughter'). In contrast, the terms for less close relatives tend to be more easily borrowed (such as the terms designating 'cousins', 'nephews', 'grandchildren', etc.). This is, by and large, the situation encountered within IE. Thus, one could argue that, over time, the original IE vocabulary dwindled, whereby 'more basic kinship terms' – so to speak – have been retained, whilst the 'less basic' ones (and others) have been replaced by borrowed terms (compare discussion below and Matras in this volume for the results of historical, contact linguistics applied to morphology).

2. 4. Is the phonological evidence malleable?

Judging from the debates outlined above, one could certainly argue that the conventional, phonological / lexical evidence on which the IE theory is based, to a closer scrutiny, appears to be rather 'malleable' – it is certainly not as decisive (one way or the other) as generally claimed. As a matter of fact, it is always possible to find a plausible, although not always testable, justification to any intervening piece of evidence counter to a stated rule or tendency. Similarly, it is also always possible to provide at least two equally plausible, equally well founded explanations for any given linguistic phenomenon. This is also the case within non-controversial areas of IE linguistics, such as the postulation of the so-called 'Indo-Iranian branch / unity'. One can in fact compare the different interpretations given to this unity – (also) on the basis of phonological /lexical evidence – in the relevant literature, for which see Lazzeroni (1998), Sims-Williams (1998) and Schmitt (1987). Compare also Schmitt in this volume, who argues that the data from both Vedic and Sanskrit are not with absolute necessity genuinely antique and that, therefore, Old Indo-Aryan is not as close to PIE as still believed by some scholars (see below).

At this point one could object that all these methodological and factual difficulties do not, after all, matter, even if they *did* impact negatively on the validity of the conventionally established sound laws (and related correspondences and reconstructions). This is because, as mentioned above, the lexicon is often considered to be the level of language less (or not at all?) relevant

for the purpose of assessing genetic relationships. This would be sound and acceptable if there were indeed a consensus among Indo-Europeanists as well as comparativists in general on the principle that it is morphology the level of language (mostly?) relevant in this context. But, do we find that consensus? This brings us to the second, specific area of debate dealt with in this volume.

3. The role of morphology

3.1. Is morphology the most reliable indicator of genetic inheritance?

Since the very times of the establishment of IE the prevailing opinion appears to have been that grammar can offer the most reliable evidence for assessing genetic relatedness. Grammar has typically been considered to be a rather stable level of language and, often, totally resistant to borrowing – in contrast to the volatility of the lexicon. These properties have made morphological correlations quite popular among many historical linguists, even if it has always been (more or less openly) recognized that it is not clear what the criteria are, if any, on the basis of which to identify and evaluate morphological similarities. In fact, not even the regularity principle is expected to consistently operate at this level, due to the overwhelming interference of the analogical principle. In addition, there has never been (to our knowledge) any systematic attempt to define a possible measure, a 'unit of similarity' (in form and / or function) – as it were – to be applied in the practice of comparing morphemes. This measure of similarity would work as a common denominator against which to evaluate the at times vaguely similar, at times very similar and at times identical morphemes occurring among (most / some) IE languages. Thus, the problem is that the morphological correlations are typically observed and established by intuitive, visual inspection (often by single scholars), whereby considerable latitude may be allowed when it comes to phonological forms as well as to similarity of functions. For the morphological (and morpho-phonological) correlations to be rigorous, to be proper

'correspondences', one should certainly require regular phonological correspondences between morphs which also indicate similar (but how similar?) meanings and /or functions – condition which is hardly ever met. Indeed, one often reads in the literature that the grammatical correlations within IE are (still nowadays) simply and purely 'obvious' to the 'naked eyes' of the trained philologist (see Nichols 1996a), exactly as they appeared to the first scholars who dealt with them a couple of centuries ago.

3. 2. The degree of 'borrowability' of grammar

In recent years a mounting body of evidence has been accumulating according to which not only grammar is found to be 'borrowable', but, given the appropriate historical and social context, it may rate quite high on the scale of borrowability. It could therefore be difficult to determine whether shared grammatical innovations are the result of genetic inheritance or of areal convergence. For example, in connection with some Australian linguistic areas the following question has often been posed: is grammar borrowable to such an extent that historical reconstruction becomes impossible? (See Dench (2001); Austin (1989); Austin (ed. 1990) & Dixon [2002]).

The issue of the (degree of) borrowability of grammar is addressed in this volume by Matras, through his detailed investigation of the grammatical borrowing that has taken place in the 'mixed' languages: Romani and Domari (see also discussion above (par. **2.3.**) with regard to their lexical borrowing). In particular, the Author shows that in these languages grammatical borrowing has been extensive, occurring even in the domain of bound morphology. Nevertheless, genetically inherited morphological elements and patterns *can* be identified and kept distinct from the borrowed ones. In other words, even in the context of extreme contact and mixing, borrowing does not appear to be random but "tends to be structured in a hierarchical manner, and so it is at least to some extent predictable". This in turn means that "some components of grammar offer more reliable indicators than others about shared historical-genetic inheritance. For example, Romani shows "unique morphological conservativism in

its nominal and verbal inflection", as illustrated by the fact that it preserves to a considerable degree the original endings of the Old /Middle Indo-Aryan present tense conjugation[7]. On the other hand, Matras also presents evidence that nominal case endings and verbal inflections, if borrowed at all, tend to be borrowed "wholesale" (see also Levin (1995), who presents evidence of this sort in the borrowing of verbal and nominal paradigms from IE into Semitic). These findings can certainly assist linguists in the delicate task of identifying original morphemes / morphological paradigms and reconstructing the morphological structure of proto-languages in general, and of the IE proto-language in particular. As a matter of fact, within IE studies there is still an open question regarding the nature of the original morphological structure of PIE: was PIE rich in morphology (as is the case, mainly, of Greek and Indo-Iranian), which has then been 'reduced' or 'lost' in the other languages, or was it rather poor in morphology, in which case the complex morphology of Indo-Iranian and Greek is the result of parallel, shared innovations, rather than of genetic inheritance? This issue is dealt with in this volume by Drinka, Kazanas and Di Giovine, who all hold different, at times contradictory views on the topic, as discussed in more details below.

This already intricate debate is further complicated by the connected issue of the 'Indo-Hittite theory' as proposed by Sturtevant (1933) – although Sturtevant's position as such has now been abandoned. In fact, although scholars appear to agree on the fact that the Anatolian languages have a special status among the IE languages, they disagree on how to interpret and justify this special status, that is, the numerous acknowledged differences existing between the Anatolian languages and all the rest of the IE languages, (particularly) at the morphological level[8]. Regarding this issue, see Carruba (1997 & 2001a[9]), Luraghi (1998); Zeilfelder (1994/2001); Marazzi et al. (1990), Drews (ed. 2001)[10], Clackson (2007: 129 ff.), etc.). For a possible interpretation of the status of Hittite within the family compare the chapter by Di Giovine, who addresses the issue of the verbal morphology, and Carruba, who addresses the peculiarities of the Hittite *Ablaut*.

Still on the issue of the 'borrowability' of grammar, it has been claimed at times that the IE morphological correlations are, on the whole, similar enough to be considered valid correlations but different enough so as not to raise the suspicion that borrowing might have been involved. However, certain IE grammatical forms typically reported in textbooks as 'obvious' examples of genetic inheritance – such as the paradigm of the verb 'to be', or 'to bear, carry' – are so similar, if not in some forms identical across the area, that the suspicion of borrowing may indeed arise. In fact, one would normally expect much more divergence from a long process of inheritance and development. These observations have been made, for example, by Croft (2005: XIX) and Greenberg (2005), and are discussed in more details in the next paragraph. In addition to this, one should take into consideration the numerous morphological correlations which supposedly connect the IE family with other contiguous, but different language families, as argued for by the supporters of the so-called macro-families (see for example Greenberg (1987 & 2005) and Comrie [1998]). Furthermore, it has recently been observed that several grammatical endings (case and personal endings), at first believed to be exclusive to IE or the Eurasiatic macro-family, actually occur also in non-contiguous areas, such the Americas, so that both the assumption of 'long distance' genetic inheritance and that of 'borrowing' have to be excluded. In other words, in addition to the traditional dichotomy: 'inherited *vs* borrowed', there is now evidence that there may also be a third, (a priori) equally possible explanation for the correlations individuated among languages at the morphological level (as well as at the phonological / lexical level): 'chance similarities', as clearly illustrated in Campbell (1995 and 2003a)[11].

Finally, it is worth reflecting on the following remarks made by Harrison (2003:223 ff.), who sums up some of the criticisms raised against the privileged role often attributed to the grammatical correlations (see also Koch (1996: 218-20) & (2001) for similar concerns):

> Grammatical objects fare poorly as evidence for genetic relatedness [...]. There can be no regularity assumption for grammatical objects to provide a measure of similarity, because grammatical objects are unique. [...]. Cases like those of morphological person-number paradigms are of particular interest because, although not universal in any absolute sense [...] linguists are surprised neither by their occurrence nor by their non-occurrence, in the verb or noun morphology, not even in closely related languages, or by the occurrence of similar paradigms in non-related languages, or *vice versa*

3. 3. Is the morphological evidence malleable?

At this point one could object that the risk of reconstructing false matches within IE grammar is rather low, since one can rely on a wealth of shared (inflectional and derivational) morphology in a great variety of areas. Furthermore, although much of the grammatical evidence put forward by traditional IE studies is certainly rather intuitive and subjective, as it happens, this evidence turns out to fall within the range of what Nichols (1996a: 49 & 64) calls "diagnostic evidence":

> Traditionally linguistic kinship was identified on the basis of diagnostic evidence which is grammatical and combines structural paradigmaticity [...] and syntagmaticity with concrete morphological forms. The Indo-Europeanists' *intuitive feel* for what was diagnostic evidence of relatedness corresponds to a computable threshold of probability of occurrence [...]. A grouping can be regarded as established by the comparative method if and only if it rests on individual-identifying evidence

The idea is that one can compute the probability of occurrence in the languages of the world of certain features, certain patterns, such as consonant sequences (in the dimension of syntagmaticity), or the co-occurrence of some grammatical categories and their morphological indicators (in the dimension of paradigmaticity), etc., whereby the Author assumes that a probability of occurrence of a given phenomenon of 1 in 100.000 is "individual-identifying". Within IE one finds several, arguably a statistically significant number of instances of individual-identifying diagnostic evidence. The following data, as proposed by Nichols herself (1996a: 47),

are among the most quoted ones for the purpose of this discussion (notice that to the Latin and Greek endings reported in this table at least the corresponding Sanskrit endings should be added to complete and strengthen the picture: M. *-as*, *-am*; F. *-ā*, *-ām*; N. *-am*, *-a*):

Table (I)

	Masculine	*Feminine*	*Neuter*
Latin			
Nominative	*-us*	*-a*	*-um*
Accusative	*-um*	*-am*	*-um*
Greek			
Nominative	*-os*	(*)-*ā*	*-on*
Accusative	*-on*	(*)-*ān*	*-on*

Here there are two dimensions of paradigmaticity: case and gender on the one hand, with 'cognates' endings combined together with identical distribution on the other. These dimensions of paradigmaticity are also quite productive: there are for example numerous items in the masculine declension: Skt. *-as* *vs* Gk. *-os* *vs* Lat. *-us* (and not just nouns but also adjectives and participles: compare Lat. *amatus*, etc.). The same holds true, for example, for the feminine declension, although other classes of nouns may be less numerous, coming down to 5-6 items only. In addition, in many cases these morphological paradigms also occur in connection with lexical items (verbs and nouns) that are typically quite similar, and are in fact considered to be cognates (see for example the various IE terms and related paradigms for 'name': Skt. *nāma-*, Lat. *nōmen*[12] etc., as well as the various cognates and paradigms of the verb 'to be' and 'to bear, carry', etc.). Several, striking examples of diagnostic evidence of this type are reported and discussed in this volume by Kazanas, who, in line with mainstream doctrine, argues that these highly complex, and at times unique patterns of lexical and morphological correspondences cannot be explained either through chance or through borrowing. This is certainly a plausible model of

explanation; however, alternative explanations are also possible. First of all (as mentioned in the previous paragraph) there is plenty of evidence that wholesale (nominal and verbal) paradigms of the type under discussion *can* also be the effect of borrowing, as shown by Matras' contribution to this volume. Second, not all the IE languages do possess those rich, (more or less) consistent, paradigmatic morphological systems we would need to establish a wealth of 'individual-identifying' grammatical evidence. In fact, this kind of evidence is found mainly among Indo-Iranian and Greek (the so-called *Greco-Aryan* model), and, to a much lesser extent, in Latin, Germanic, Balto-Slavic, Armenian and Celtic. As to Hittite[13] and the other Anatolian languages, there appears to be consensus on the fact that the "Greco-Aryan model does not work well as an explanation for the Anatolian verb" (to use Clackson's (2007:115) words). As mentioned above, many scholars explain this situation by assuming that the less morphologically rich languages have 'lost' / 'reduced' the original, complex morphological / paradigmatic systems present in the proto-language[14] (this model is in line with the original views by Bopp, Schleicher, Brugmann, etc., who reconstructed indeed a nominal and verbal system very similar to that of Greek and, particularly, of Sanskrit). In this volume this view is embraced by Kazanas as well as (with some reservations) by Di Giovine, who also puts forwards several proposals on how to account for the peculiarities found in the Hittite verbal system. However, other scholars think that the opposite holds[15], that is: the special bond existing between Greek and Indo-Iranian in their morphological (particularly verbal) system appears to point (mainly) to widespread processes of possibly late contacts and areal diffusion[16], rather than to archaism. This position is defended in this volume by Drinka, and, to a certain extent, implicitly, by Schmitt, when he claims that (Vedic) Sanskrit is not necessarily the oldest language within IE (see below). In particular, Drinka proposes a "stratified model of PIE, one which takes into account the systematic, layered morphological correspondences of Greek and Indo-Iranian". This in turn means that those "languages which share only archaic elements, such as Hittite and Germanic, are presumed to have

separated from other IE languages at an earlier time". Thus, Drinka relies on the "stratification of data" – innovative layers building on more archaic layer – to account for the morphological correlations observed among the IE languages (for a 'stratified' interpretation of IE see also Andersen's chapter).

The idea of the existence of some sort of 'Indo-Greek branch / unity' is not totally new, having already been proposed, for example, by Pisani (1933, 1940 & 1974). As is known, this idea was mostly rejected at the time, since there are several other grammatical (as well as phonological) isoglosses that cut across such an assumed branch (compare for example, the reduplicated perfect of Germanic, or the future ending in *-s* of Lithuanian, shared with Sanskrit and Greek[17]). Some scholars still nowadays continue to reject the thesis of the Indo-Greek unity, arguing that it is a rather naïve notion, based on selected, partial evidence. This is clearly stated in Kazanas (2004) and claimed by Kazanas in this volume. However, Drinka (in her contribution to the volume), on the basis of new data and new methods of analysis, argues that this unity does indeed exist, and uses the definition: "eastern languages" to refer to the Greek and Indo-Iranian special morphological bond. On the other hand, Di Giovine (in this volume) argues that looking at the issue in simple, binary terms of 'archaism' *vs* 'innovation' does not lead anywhere, since the status of the IE verbal system is much more intricate and subtle than this dichotomy would lead us to believe. All this, once again, ties in with, and lends support to, the well known fact that a clear-cut IE family tree is still difficult to draw, both at the phonological and morphological level[18]. This is why Drinka insists that areal / contact models of interpretation have to be adopted in addition to, or, actually, incorporated into, the traditional paradigm: only an integrated, three dimensional model ("an amalgamation of family tree and wave models") can "explain" the nature and distribution of the correlations found among the IE languages (as also pointed by Andersen (1998: 420) and in this volume).

Whatever the case, one has to admit and reflect on the fact that the morphological reconstructions are essentially based on a few core languages, whose morphological correlations – whether

the result of archaism, or innovation, or borrowing, or convergence – are nevertheless well attested, whilst "the attribution of the other languages to the family is necessarily done on partial evidence" (as stated by Morpurgo Davies; Cambridge Seminar[19] [2005]).

3. 4. Are morphological reconstructions predictive?

It is reasonable to ask at this point whether the circularity issue (as discussed above with regard to phonology and lexicon) could have a negative impact also on the morphological reconstructions, including the area of diagnostic evidence. In fact, here too scholars can increase the power of the comparative method by making recourse to a number of 'adjustable parameters'. Actually, the risk of circularity is higher at the morphological level, because of the (supposed) pervasive operation of analogy and the lack of a rigorous definition of 'morphological correlation / similarity' (as mentioned above). One can compare at this regard the many difficulties, obscurities and exceptions encountered in the reconstruction of the IE verbal and nominal paradigms. For example, with regard to the (thematic) *o*-stem[20] class of nouns referred to in Table (I) above (certainly a highly productive class of nouns, displaying rather consistent paradigmatic patterns), one can nevertheless find several, more or less serious difficulties. For example, some case endings cannot be reconstructed in an economic, straightforward way, or cannot be reconstructed at all, so that a number of 'adjustable parameters' have to be introduced in order to overcome the observed mismatches. The adjustable parameters in question include:

1. a chain of (at times unverifiable) assumptions, including syncretism and reshaping through analogy;
2. a chain of (at times unverifiable) intermediate reconstructions and alternating forms;
3. a chain of minor but frequent (and, at time unverifiable) language-specific sound changes,
4. if any of these procedures fail, there is always the possibility of giving different interpretations to the internal structure, function and origin of the case endings themselves.

As concrete examples of the complications involved in the reconstruction of several case endings in this class of nouns compare the following[21]:

1. The instrumental singular, whose reconstruction requires the help of a laryngeal segment and vowel alternation: **-e-H_1* ~ **-o-H_1*, whereby the ending is unmistakable attested only in Indo-Iranian *-ā* [22] (although here too only in a dozen or so forms in the Rigveda, the usual ending being, even in the Rigveda, *-enā*); no *o*-stem form is quotable from Mycenaean.
2. The genitive singular, whose overall picture is not clear, and for which at least two alternative reconstructions have to be postulated: a) the alternating type **-os(y)o* ~ **-es(y)o*, attested in Indo-Iranian, earliest Greek and Italic; possibly also in Germanic and Balto-Slavic, although with some obscurities in its development, and, possibly, in the Hieroglyphic Hittite *-asi*; b) the type **-ī* (Italic), found in Latin and Celtic (whereby scholars appear to be still divided as to whether the two types of endings do or do not have an ultimate, common origin; see also Hock [1986:453]).
3. The locative singular **-ey* ~ **-oy*; here there are some doubts as to which of the two variants may have been the original form.
4. The ablative singular, variously reconstructed as **-ad* ~ **-at* ~ **-$h_2$et* (whence *-āt* in Vedic and **-ōt* in Italic). This does not appear to have the expected reflex in Lithuanian and does have some other, minor oddities in Latin adverbs in *-ē*.
5. The nominative plural **-ōs* in nouns and **-ōy* in pronouns, whose original distribution (Indo-Iranian, Germanic, Osco-Umbrian, and, with a bit of remodeling, Hittite [*at-te-eš*]), was otherwise "disturbed" (to use Sihler's (1995:261)

words), as shown in Old Latin *poploe*. In the Rigveda fully a third of the endings of nominative plural are represented by another ending: *-āsas*, which also occurs in Avestan.

6. The dative / ablative / instrumental / locative plural, which are expressed by various endings: a)*-ōys* (attested in Vedic, Greek and Italic; see Old Latin *poplois*); b) *-ebhis*, which is the instrumental plural, occurring in Vedic with almost the same frequency as the ending *-āis*, derived from *-ōys* (it is questionable whether an ending of instrumental case can be equated to that of a dative/ablative/locative); c) Old Indian *-bhyas*, to which Latin *-bus* in the other declensions corresponds; however, in Germanic and Balto-Slavic there is an ending *-m-* (*-mis*); d) locative *-oi-su*, present in Vedic, Homeric Greek, etc.
7. Finally, notice the Genitive plural *-ē* of Gothic, "still an unsolved problem" (according to Szemerényi [1996:185]).

Accounting for these morphological difficulties (however minor they may look) requires a number of justifications which all contribute to build up the complexity and, most likely, the 'post-dictive' nature of the explanatory apparatus.

3. 5. Counter evidence within morphology

It has been argued above that the morphological evidence on which the IE theory is founded appears to be rather 'malleable', this problem being caused by the general lack of criteria and guidelines on how to evaluate morphological correlations. This in turn means that it is equally difficult, if not impossible, to identify evidence potentially counter to the model (with perhaps the only exception of the evidence from languages in contact that wholesale morphological paradigms *can* be borrowed). For example, the absence of the fundamental IE category of feminine gender in Anatolian[23] could be considered a paradigmatic example of the inability on behalf of the IE theory (and the methods of historical linguistics in general) to make very clear-cut, testable predictions. There have been no claims (as far as we know) that the absence of this 'diagnostic' category should be classified and set apart as

potential evidence counter to the IE origin of Anatolian[24]. Quite the contrary, efforts have been devoted only to finding traces of the feminine gender here and there or to justify its absence in the most plausible, natural way (see for example Oettinger (1987); Melchert (1992); Justus[25] (2002), Clackson (2007:104ff.), etc). The same could be argued with regard to one of the most puzzling aspects of Hittite verbal morphology, the 'ḫi-conjugation', which, admittedly, does not easily slot into any reconstructed category of IE (although there are claimed to be some good lexical correspondences between verbs adopting this conjugation and verbal roots in other IE languages; see Jasanoff (2003); Neu & Meid (eds, 1979); Clackson (2007:139 ff.) and Carruba in this volume). The situation does not appear to be much different at the level of morpho-phonology, as is evident from the state of the research regarding the IE *Ablaut* (vowel alternation / apophony). Here, the scanty and irregular distribution of vowel alternations across the IE area has generated an array of complex, sophisticated explanations, none of which however appears to be either testable or satisfactory, as clearly illustrated in the chapter by Carruba in this volume. Carruba suggests that the original function of *Ablaut* is to be sought in the 'deixis' – a rather archaic, elementary function, but a very productive one within IE. He ascribes the poor attestation of coherent patterns of vowel alternations across the IE area to the very ancient origin of phenomenon itself. This is certainly a plausible explanation; however, again, there may be alternative explanations. For example, it could be argued that the *Ablaut* present in Sanskrit (and partially in Greek), does not represent a phenomenon to be traced back to the IE proto-language, but on the contrary, a language-specific phenomenon. This is, in fact, basically Kazanas' opinion, as expressed in his assessment of *Ablaut* in this volume: Sanskrit would be the only IE language that has roots and a "proper vowel gradation". Compare also Marcantonio's contribution, which investigates the 'statistical significance' (or lack of it) of the present *vs* perfect vowel alternation. Finally, compare Di Giovine (1995) and Di Giovine in this volume, who instead embraces a more conventional approach to IE *Ablaut*.

3. 6. Is there hard core evidence at the foundation of Indo-European?

Bearing in mind the shortcomings of the morphological correlations as discussed above, it is understandable how several scholars, such as Campbell (2003a & b) and Morpurgo Davies (Cambridge Seminar 2005; see note [19]) have reiterated and emphasized the role of the phonological (/lexical) correspondences for the purpose of assessing genetic relations. One could compare also the case of Proto-Dravidian, where it is the lexicon that has seen the greatest amount of relevant, comparative work, as illustrated by Steever (1996). As a matter of fact, phonology has at times offered us the possibility of testing the validity of the comparative method, in the sense that some predictions made by the comparative analysis have subsequently been proven correct in connection with the discovery of new linguistic material. For example, new evidence from Hittite and Linear B Mycenaean has confirmed the validity of earlier reconstructions, in this way also allowing us to redefine the processes which lead to the attested languages. The evidence in question includes:

1. The laryngeal segments found in Hittite, which would fill in and confirm the existence of the 'coefficient sonantique' predicted by Saussure in his *Mémoire*.
2. The so-called '*r/n*-stems' (the best known class of 'heteroclitic' stems), an archaic group of nouns[26] displaying a different final consonant in the direct *vs* oblique cases. This alternating stem had been reconstructed (through the means of both internal and external comparison) long before further evidence for the alternation became available with the discovery of Hittite, where, in fact, this class of nouns is somewhat productive (contrary to what happens in the other languages, where it is a relict; compare Latin *iter* ~ *itineris*, or Greek *ὕδωρ* ~ *ὕδατος*).

3. A series of separate sounds in Mycenaean, testifying in favor of the postulated IE labio-velars stops, which were lost by the time of alphabetic Greek (having merged with other stops according to specific rules and dialects).This distinct series of stops, transliterated as *q*, not only are attested in Linear B, but are also still preserved exactly where predicted[27], as shown, for example, by *qe* 'and', with *q* as a reflex of labio-velar (< $*k^we$).

These are undoubtedly remarkable discoveries, and fine analyses. However, the data and analyses under discussion are not void of difficulties. For example, with regard to the Mycenaean material and the postulation of labio-velar segments, one should bear in mind that, given the Mycenaean spelling system, we do not really have direct proof that the segment in question is indeed a labio-velar, although this represents, without doubt, "a strong contender" (to use McMahon & McMahon's (2005:14) words[28]). As to the heteroclitic stems, it is recognized that the Hittite evidence is contradictory, or at least unsatisfactory, since not all the *r/n*-stem Hittite nouns available appear to have IE counterparts, as clearly pointed out by Sihler (1995: 299-300). Furthermore, one could observe that the more linguistic material, the more *comparanda* one brings into the equation whilst doing comparison, the easier it is to find a phonological /lexical (as well as morphological) match with a given form or reconstruction. As Campbell (1998: 277) puts it, referring to the issue of mass-comparison and macro-families: "The potential for accidental matching increases dramatically ... when one increases the pool of words from which potential cognates are sought or when one permits the semantics of compared forms to vary even slightly" (for similar remarks see Hock (1993)[29]; compare also the chapter by Marcantonio in this volume). Finally, one could always argue that a few instances of fulfilled predictions – the occurrence of the right, predicted reflex at the right place – are not enough to prove the supposed predictive nature of all the conventionally established IE laws.

As to the laryngeal theory, it is undeniable that it represents one of the most outstanding examples of a successful theory within

historical linguistics. As Andersen (2007)[30] puts it: "From a purely algebraic theory in Saussure; given putative phonetic justification by Hermann Moeller; then actually justified by the discovery of some of the posited segments in Anatolian; found to correlate with the 'prothetic[31] vowels' and 'Attic reduplication' in Greek, and of 'final lengthening' in Homer; and then with the Baltic and Slavic distinction between acute and circumflex long vowels and diphthongs". Nevertheless, the laryngeal theory is not uncontroversial, at least with regard to the number of laryngeal segments to be postulated. Compare for example the divergent views by Martinet (1953), Szemerényi (1973 & 1996:121-6), Meier-Brügger et al. (2003:107), Lindeman (1982), etc. Compare also, for example, Beekes (1989:1 ff.), who argues in favor of three laryngeals, whose number would be required by Greek, although Hittite "certainly points to two laryngeals", and Di Giovine (2005), who, in line with other scholars, calls into question the postulation of $/h_1/$, since this segment does not have direct reflexes in any of the IE languages, including Anatolian[32]. However, the main problem associated with the laryngeal is that the advantages attained by making recourse to the laryngeal theory are counterbalanced by a number of disadvantages. For example, Winter (1990: 20-1) referring to Saussure's original idea of the 'coefficient sonantique' and subsequent developments, describes one of the known 'disadvantages' of the laryngeal theory as follows:

> We attempt to simplify our statements by subsuming overtly differing phenomena under one common formula which may or may not require positing directly unattested elements conditioning the differences [....]. This analysis had the tremendous advantage that now all subtypes of ablaut in Proto-Indo-European could be treated alike and that canonical forms of roots could be set up; along with this, however, went the disadvantage that the Proto-Indo-European system of vowels apparently had to be reduced in an unreasonable way

Sihler (1995:111 ff.), who reconstructs an *Ablaut* system "considerably leaner" than others thanks to the support of the

laryngeal theory, appears to (implicitly at least) identify another, major 'disadvantage'[33]. In fact, he states that the economy of his model is achieved by "a complication elsewhere in the system", because his reconstruction also requires "a number of sound laws" applying to the postulated laryngeals[34]. Here, the fundamental questions (which have not been asked yet, to our knowledge) are: a) how many more sound laws do we need to make the laryngeal system work? b): is this extra number of segments and required operating rules added to the system as high as to nullify the otherwise attained benefits? In this regard, Marcantonio (in this volume) argues that adding laryngeal segments to reconstructions does indeed increase the explanatory power of the comparative method[35].

Thus, if we accept the analyses, objections and counter objections expounded thus far, it is till not clear what exactly the uncontroversial, hard core evidence that lies at the foundation of the IE theory is supposed to be[36].

4. Is there a pre-historical reality behind reconstructed Indo-European?

4. 1. Conventionalism *vs* realism

The 'conventionalist' *vs* 'realist' approach to linguistic reconstruction, and related concept of proto-language (in general as well as within IE), is another long standing debate within historical linguistics, and one for which, yet again, there does not appear to be much of a consensus. Compare Koerner (1989) for an overview of the debate; compare also Bonfante (1945); Pulgram (1959 & 1961); Meid[37] (1975) and Clackson (2007:16 ff.). It is true that between the extreme conventionalist approach on the one side and the full realistic approach on the other there are plenty of more moderate, intermediate positions. Nevertheless, it appears that the fully realistic approach, which in turn is based on the controversial method of palaeo-linguistics[38], is the one that has so far attracted many supporters, and not only among archaeologists (such as Mallory (1989 & 2001) and Renfrew [1987, 1990]), but also among linguists. For example, in his criticism to the thesis by

Gimbutas (1970) regarding the location of the IE home land, Schmitt[39] (1974: 283) observes that "[with] the methods of linguistic paleontology anything can be proved as Proto-Indo-European, but it can *not* be proved as *typically* Proto-Indo-European". Thus, Schmitt appears to have confidence in the validity of the realist approach, although he has doubts regarding the factual interpretation of single reconstructions and the hasty conclusions reached about the reconstruction of single 'pre-historical facts'. It could be said that in this case too (like in the case of the 'regularity principle') it is not always made explicit what the motivations are, what the evidence is, that provides the basis for the choice of one approach, or the other. In this volume, this issue is dealt with, or touched upon, by Andersen, Bryant, Drinka and Di Giovine. The chapter by Andersen deals (also) with the ontological status of proto-languages and the relation between such a 'language' and community languages. The chapter by Bryant (also) discusses the 'malleability' of the palaeo-linguistic evidence relating to the origin of the Sanskrit language and culture. The chapter by Drinka draws attention (among others) to the fact that the linguistic evidence (in particular, specific morphological correlations) provides more complete information as to the relatedness of population and to historical /cultural reconstruction than archaeological records. Di Giovine, on the basis of the analysis of the IE verb inflectional system presented in his contribution, argues that PIE was not a compact language or not even a community language at all. Compare also the chapter by Häusler, as discussed in the following paragraph.

4. 2. The Aryan debate

Within the camp of the realist approach the debate has revolved mainly around the issue of the whereabouts of the (assumed) IE proto-community. A strand of this debate has recently re-surfaced under the definition of: 'the Aryan Debate', for which see Bryant (2001), Bryant & Patton (eds, 2005), Kazanas (2001, 2002, 2003 & 2004); Polomé (1985); Tripathi (ed., 2002/2005); Masica (1991:32-60), as well as the two debate-volumes on the origin of complex societies in Central Eurasia, edited by Jones-Bley & Zdanovich (2002). On this topic, compare the chapter by Bryant, Drinka, Häusler and Kazanas (and, to a certain extent, Schmitt) in this volume. The Aryan debate centers on the following, major interconnected issues:

1. Was the assumed homeland located in the west, somewhere in Europe, or in the east, somewhere in North India /Pakistan? In other words, were the bearers of the Old Indo-Aryan (Vedic and Sanskrit) culture indigenous or intrusive in North India?
2. Why, how, when and in which direction did the original proto-community disintegrate and migrate / spread, so as to bring about the distribution of the IE languages as we found them in historical times?
3. How old is Old Indo-Aryan? Is it as old as the traditional IE theory claims (about 1700/1500 BC), in line with the 'migrationist' model, or is it much older than that, in line with the 'indigenist' model? In turn, these questions are connected with the following major issue (already touched upon above):
4. Is Old Indo-Aryan the most archaic, and therefore, the most important language for the purpose of reconstruction?

The issues raised in points (3) and (4) are of particular relevance, not only because they are strictly connected to the home land issue, but also because the assumption of the great antiquity of (Vedic) Sanskrit is (or, at least, it has been) one of the pillar tenets at the foundation of the IE theory, and related traditional reconstructions. As we have seen above (par. **3. 3.**), this tenet is now under attack, since several scholars believe that the *Greco-Aryan* morphological paradigm represents in fact innovation, and not archaism. As to the more specific issue of the archaicity of Sanskrit, Schmitt in this volume calls into question this traditional view, arguing that: "the archaism particularly of the Old Avestan language makes it only too clear that, despite the old age of the earliest Vedic texts, the Old Indo-Aryan language is not the only fundament of the IE proto-language [...] and that its data are not

with absolute necessity genuinely antique; therefore Old Indo-Aryan (both Vedic and Sanskrit) is not so close to PIE as many people think". On the other hand, Kazanas (quoted, and in this volume) supports the traditional way of thinking as well as the indigenist model claiming exactly the opposite: Old-Indo-Aryan is indeed the oldest language within the family (actually, much older than conventionally assumed), as well as the closest language to PIE, even though a PIE language cannot be reconstructed. In turn Drinka, in her contribution to this volume, rejects the indigenist / 'Out-of -India' model, arguing against it (mainly) on the basis of those very "archaic" morphological isoglosses Indo-Aryan shares with some IE languages and those very "innovative features" it shares instead only with Iranian and Greek.

4. 3. Sanskrit and the South Asia linguistic area

This already articulated debate is in turn connected with the other, equally tangled issue of how to interpret the lexical, phonological, structural and even morphological correlations that have been identified as existing between the Old (and Modern) Indo-Aryan languages and the other non IE languages of India, such as (mainly) Dravidian and Munda. In fact, these (and other) various languages / language families of India are widely claimed to form what is usually referred to as the 'South Asia linguistic area', for which see Steever (1993:10 & 1996:11 ff.); Masica (1976) and Emeneau (1956, 1971 & 1980). In particular, on the basis of (supposed) lexical borrowing from Dravidian into Indo-Aryan and from the modern geographical distribution of the Dravidian languages, it is often claimed that Dravidian and Indo-Aryan must have been in contact – since prehistoric times– in those extreme northwestern areas of India first inhabited by the Indo-Aryans (but see Hock (1996:38) and discussion below for different points of view[40]). Thus, the relevant question here, the question which has been hotly debated and whose resolution, if ever attained, could help to break the deadlock of the indigenist *vs* migrationist model – together with the issue of the degree of antiquity of Sanskrit – is the following: Are the non IE features present in Old-Indo-Aryan the result of sub-stratum, super-stratum, ad-stratum, convergence, or even genetic inheritance? At this regard, compare the (at times clashing) views by scholars such as: Burrow (1946, 1973 & 1975); Emeneau (1980); Hock (1975, 1986, 1996 & 1999); Kuiper (1955 & 1991); Witzel (1999 & 2000); Kuzmina (2003); Misra (2005); Kazanas (2004) and Kazanas in this volume. Several scholars, such as Southworth (1979), Kuiper (1955, 1991 & 1992) and Burrow (1973), have compiled long lists of (Rig-)Vedic words believed to be of Dravidian or, more generally of non-Aryan origin. Other scholars, particularly Hock (1996: 36ff.), have drawn attention to the fact that the situation is not that clear-cut after all. In fact, in many instances it is difficult to trace back the origin of the lexical as well as structural similarities observed in this intricate linguistic area, and to sort out whether they are attributable to borrowing, convergence, or chance resemblances[41]. On the other hand, there are Indian scholars (such as Misra (2005) and Aiyar [1975]) who challenge the traditional classification of Sanskrit as an IE language, observing that under a different socio-politico-cultural context, the similarities under discussion could have been interpreted differently. For example, they could have been ascribed to the postulation of a shared, common ancestor for Sanskrit and Dravidian, and the studies of the South Asian linguistic area, as we know them today, might have taken a totally different course, generating a different production of knowledge, and, possibly, a totally different proto-language as the ancestor of Sanskrit. In turn, these scholars are accused to be strongly biased in their research, being driven by nationalistic feelings and political motivations. Certainly – it has to be said – there are elements of truth in this statement, since the main concerns (mostly unexpressed) of these Indian scholars are for the cultural unity of India under a Sanskritic proto-banner, whereby Dravidian would be some sort of Sanskritic prakrit (on these issues see the chapter by Bryant and Annamalai & Steever in this volume). Nevertheless, it is also fair to recognize that had the events, the accidents of History been different, in other words, had Sanskrit not been brought to the attention of the western scholars (as a consequence of the English colonization of India, for which see Bryant [2001]), the cultural scenario of historical linguistics might

well have been totally different indeed, and one can easily imagine how a different paradigm might have been created to account for the origin of Sanskrit and the other languages (/language families) of India. Actually, the fact that the initial choice of comparing a certain pair / set of languages, rather than others, is by its own nature 'intuitive' and may be also influenced[42], or even dictated by circumstances external to linguistics, can increase the risk of circularity embedded in the traditional methods of analysis, as discussed above (for similar remarks see Clackson[43] [2007: 3]). It would therefore be desirable to be able to 'demonstrate' that the correlations shared by Sanskrit and Dravidian, for example, are the genuine effect of a typical process of *Sprachbund* convergence, rather than simply 'labeling' them as 'contact-induced' just because it is widely accepted that the languages in question are not genetically related. The thesis that the correlations shared by Sanskrit and Dravidians are the genuine effect of *Sprachbund* is embraced by Annamalai & Steever. In their chapter the Authors do not directly address the issue of the origin of the non IE features present in Old (and Modern) Indo-Aryan languages, but argue that Dravidian is a clearly distinct linguistic family from Indo-Aryan, although the latter does indeed display Dravidian features. In other words, according to Annamalai & Steever, "some of the four genetically distinct language stocks in South Asia have clear genetic linguistic relations outside of the subcontinent". As a matter of fact, a comparison of two classical languages of India, Sanskrit and Old Tamil, as carried out by the Authors, would reveal and illustrate a typical *Sprachbund* situation, where "two languages may be genetically distinct, yet grammatically related".

At this point it is interesting to observe that there is at least one aspect within this intermingled debate about which there appears to be consensus among all scholars: there is no archaeological evidence for the assumed migrations of (part of the) original IE community, either from a supposed western homeland eastward into North India, or the other way round, as pointed out by Bryant (2001), Chapman & Hamerow (eds, 1997) and Häusler (2002, 2003a, b& c, & 2004). Compare also the contribution by Häusler and Kazanas in this volume. Häusler (2004:1) points out

that the explanation of the spread of the Indo-European languages as "the spread of the culture of a concrete group of people from a cradle [...] has not been proved by anthropology and archaeology as yet". The Author (*ibidem*) also believes that "all attempts to reconstruct the old culture of the Indo-Europeans as existing in a concrete cradle, by the means of 'linguistic paleontology' are wrong". On the other hand, the objection could certainly be raised that the 'absence of evidence is not evidence of absence', so that scholars are faced here with yet another area of IE studies where the 'evidence' (or lack of it) is malleable and un-decisive. As a matter of fact, Häusler himself believes that "the IE linguistic community must have existed at some point in ancient times, since the linguistic classification has been fairly safely established, even if this (assumed) linguistic community is not at all retrievable by the means of archaeological and anthropological research" (as stated in his contribution to the volume). Whatever the case, the fact remains that, thus far, the required, supporting (one way or the other) archaeological and palaeo-anthropological evidence has not been found.

As to the palaeo-linguistic evidence – place and river names, terms for flora & fauna, names of deities, etc. – mainly from Sanskrit, some scholars claim that this evidence is malleable, and therefore inconclusive when it comes to trying to establish the indigenous or intrusive character of Old Indo-Aryan; see Bryant (2001) and Bryant in this volume. In contrast, other scholars claim that the palaeo-linguistic evidence clearly supports the indigenist / 'Out-of-India' model; see for example Kazanas (2004) and Kazanas in this volume. Drinka, on the other hand, argues that the thesis of late contacts between Greek and Indo-Iranian – as put forward in her contribution to the volume – constitutes one more piece of evidence that the 'Out of India' theory is untenable.

Finally, it is worth mentioning that recent genetic evidence (whatever its relevance may be in this context) appears to support the indigenist hypothesis of the origin of the Sanskrit speaking peoples and culture; see Kazanas (2002 & 2004), as well as Kazanas' and Bryant's chapter in this volume.

4. 4. The conventionalist approach

Those scholars who adopt the conventionalist approach to reconstruction claim that linguistic classifications do not necessarily imply and guarantee the existence of the corresponding speech community. Therefore, searching for the original IE community, and the pre-historical, highly intertwined linguistic and extra-linguistic processes that would have brought about the IE languages, is a pointless task. This remains the case even if linguistics is assisted by other disciplines whose methods of analysis are claimed to have recently achieved high levels of reliability, such as archaeology and genetics. In other words, comparative historical linguistics cannot in any way shed light onto pre-historical 'facts' (homelands, migrations, institutions etc.). Historical linguistics was not meant to be, it is not and cannot be used as a 'branch of prehistory'. As Harrison forcefully (2003:231) puts it:

> Too many comparative historical linguists want to dig up Troy, linguistically speaking. They consider it more important that comparative historical linguistics shed light on prehistoric migrations than that it shed light on the nature of language change [...]. I do not consider comparative historical linguistics a branch of prehistory, and I sincerely believe that if we cared less about dates, maps, and trees, and more about language change, there'd be more real progress in the field

5. Introduction to the chapters (in alphabetical order)

Chapter II. Henning Andersen's chapter: *The* satem *languages of the Indo-European Northwest. First contacts?* focuses on some of the key concepts of comparative linguistics: *Stammbaum*, proto-language, the relation between stemmas and the '(re)construction' of language histories, diffusion, Schmidtian dialectal areas, etc. The Author's specific field of research is the Slavic and Baltic languages – a very good starting point for the debate in question. In fact, these languages show obvious similarities, underpinned by regular and systematic correspondences on the one hand, and

"significant irregularities in phonological development, divergences in morphology and deep differences in vocabulary" on the other hand. This in turn means that a coherent, pre-historical account of their origin and mutual relationships proves difficult to draw. In particular, Andersen observes that even some of the regular correspondences holding within and among the two branches, such as instances of regular developments of their respective accent (/accent system), may be deceiving, in the sense that it is difficult to ascertain whether these developments are the result of shared innovation or, instead, of parallel but independent developments, triggered by language contact (including contact with pre-IE languages). Actually, the Author argues, some of the earliest accentual features shared by Baltic and Slavic, traditionally interpreted as shared innovation and, therefore, as inherited, are in fact most likely to be the result of separate development. The Author further remarks how the laws conventionally established to account for the Balto-Slavic accentology cannot really be considered 'law', since they are merely plain, "laconic restatement" of the observed phenomena, often relying upon massive recourse to unjustified (and atypical) analogical changes. A new, more rigorous interpretation of the relevant data and processes, a "paradigm shift", is then proposed. The Balto-Slavic issue offers Andersen the opportunity to re-affirm his thesis of the "stratification of the IE Northwest", according to which the substantial differences existing between the two language groups as well as within Baltic itself can be accounted for by positing various, distinct substrata, distinct IE language traditions underlying the later pre-Baltic and pre-Slavic dialects. In other words, the European Northwest would be one of those areas where "admigrations" took place, i.e., migrations staggered at different times – just as there must have been other admigrations from the IE core area. It could be said that Andersen criticizes IE /comparative linguistics from within – as it were – accepting the basic paradigm and traditional procedures whilst focusing on appropriate and inappropriate ways of interpreting the established correspondences and tenets, in an attempt to attain realistic and more broadly acceptable results. At this regard the

reader might want to compare the chapter by Marcantonio, who calls into question the (supposed) 'scientific' nature of the comparative method and, as a consequence, of the IE family. Compare also the chapter by Drinka, who shares with Andersen the view that a "stratified model of IE relationship" can better account for the relevant evidence. Kazanas, on the other hand, lines up with both Andersen and Drinka in reaffirming the validity of the IE theory in general, but he has no faith in the conventional IE reconstructions.

Chapter III. The chapter by E. Annamalai & S. B. Steever: *Ideology, the Indian homeland hypothesis and the comparative method*, focuses on a wide range comparison of two classical languages of India: Old Tamil and Sanskrit. By paying particular attention to the phonological, morphological and morpho-phonological components of their respective grammar, the Authors point out where and how these languages converge and where and how they diverge. After evaluating the nature of the observed similarities and differences, the Authors argue that Old Tamil and Sanskrit belong to two distinct language families (Dravidian and IE, respectively), according to mainstream view. They also reaffirm the conventional thesis that the shared similarities of these languages clearly bear the diagnostic marks of a *Sprachbund* convergence. Thus, according to Annamalai & Steever, the application of the comparative method to the Indo-Aryan languages continues to uphold the traditional IE hypothesis, by means of the well known, systematic correspondences existing among the various IE languages. On the contrary, those scholars who put forward claims about the historical relation of the South Asian languages to an Indian homeland – also by calling the methods of historical linguistics into question – find themselves at a loss to economically explain the systematic correlations that hold within IE. More specifically, the proposed comparison of the selected structural features of Sanskrit and Old Tamil is claimed to show how these two languages do not incorporate reference to a so-called 'geographic index' – an index that would point toward a historical origin of the languages in question in South Asia. As a

consequence, the hypothesis of an Indian homeland, as well the criticisms raised against the comparative method, would owe more to ideological currents of thought than to scientific, linguistic analysis. At this regard, the reader may want to compare the views expounded by Annamalai & Steever with those put forward by Bryant and Kazanas. For example, Bryant points out how major areas of the conventional (linguistic and extra-linguistic) evidence on which the (Indo-)European origin of Sanskrit is based, can actually lead to a circular reading. Kazanas instead supports the Indigenist hypothesis for the origin of Sanskrit (/Indo-Iranian) and reaffirms its IE classification, but denies that the IE languages are held together by regular and systematic correspondences.

Chapter IV. Edwin Bryant, in his chapter: *The Indo-Aryan migration debate* summarizes the debate relating to the origin of the Vedic culture and language. The numerous, intricate facets of this debate include: the underlying ideological background, some of the more prominent aspects of the data brought forward by either camp (migrationists *vs* indigenists), the methodological shortcomings underpinning the relevant evidence, including a widespread, embedded circularity, etc. The main claim made by the Author is that the relevant palaeo-linguistic evidence, as well as some aspects of the extra-linguistic evidence, is so flexible, so 'malleable' that it does not permit a determination of whether the Vedic culture is of local origin or is instead intrusive into North West India. More specifically, much of the evidence associated with the Indo-Aryans, whether philological, linguistic or even archaeological (such as the absence of horse and chariot remains in the relevant archaeological sites), can be construed in sometimes diametrically opposed ways by scholars approaching the issue from differing angles. With little or nothing in the data itself to favor either models of interpretation, marshalling the data in support of one or the other position is clearly a consequence of the starting up assumptions. Then, typically, the selected reading of the data is configured to support the initial assumption, in this way generating an obvious circularity. This circularity becomes particularly acute in the case of the vexed issue of the (supposed)

borrowed elements in Sanskrit. Here again, on many instances the data itself would not allow to trace back the source and /or direction of the borrowing. Thus, if one did not accept that the IE classification is well founded, the overall borrowing issue, the connected substratum *vs* ad-stratum issue, and, ultimately, the conventional classification of the Indian languages involved (including Dravidian and Munda) would certainly demand revisiting. Bryant's views are at odds with the views of those scholars (in general and in this volume) who argue instead that there is enough linguistic and extra-linguistic evidence to allow linguists to support either the indigenist model (see Kazanas) or, conversely, the migrationist model (see Drinka, Annamalai & Steever and Schmitt). Bryant's chapter can be read also in connection with the chapter by Marcantonio, who challenges the normative model of IE historical reconstruction. Finally, Bryant points out that there is at least one clear, undeniable 'fact' amongst so much malleable evidence: no trace of migration (in either direction) can be found in the field of archaeological, paleo-anthropological and historical research (as also claimed by Häusler), whilst the genetic evidence appears to suggest a local origin for the bearers of the Vedic culture (as also pointed out by Kazanas).

Chapter V. Onofrio Carruba, in his chapter: *Indo-European vowel alternations: (Ablaut/ apophony)* intends first to review the state of the research regarding the IE vowel alternations, and then to put forward his own views regarding the origin and nature of this phenomenon. In fact, the Author argues, despite intense research and long standing debates, neither the origin nor the original function of the apophony have really been identified. More specifically, Carruba intends to investigate the connection between the "mobility" of position of the accent and the alternating vowels – and related "mobility of meaning" – both at the PIE and the "pre-PIE" stage of development. To do so, one should depart from the traditional schemes of analysis and investigate instead the relationship between the stem of the present in *-e* and that of the perfect in *-o vs* the vowels in question. It is here in fact that one

can clearly observe the 'deictic' value and origin of the apophony, both in the sense of its 'temporal' value ('vicinity' (/presence) *vs* 'distance'), and in the sense of its 'markedness' value (*-o* being marked and *-e* unmarked). This would mean that the accent – now such a vital element for the apophony – must not have been that vital, or must have been a pitch type of accent, in an earlier phase of IE. To trace back the ablauting structure of this pre-PIE phase it is not only necessary to take into account the morpho-phonemic schemes of the actual IE languages, but also to make recourse to linguistic notions, linguistic elements that are proven to be really archaic, as is indeed the 'deixis' (a rather elementary notion, but a very productive one within IE). Having identified the 'deixis' as the founding factor of the (pre-proto-)IE apophony, Carruba focuses his attention onto the Anatolian languages. Here we would observe a clear example of the "breaking down" of the typical IE ablauting schemes (that are in any way well preserved only in Old Indian, and, to a lesser extent, in Greek). In fact, Hittite typically displays several new ablauting forms (such as *sāk(ah)hi* 'to know'), although it also preserves some old, known ones. Carruba concludes that: "There are still numerous difficulties that hinder scholars from fully understanding the nature and behavior of the apophony [....]. It appears to be lost in the depth of time, at the point that it is practically impossible to achieve an un-equivocal and coherent reconstruction". In line with mainstream doctrine, Carruba ascribes these difficulties of reconstruction to the 'loss' of the relevant *comparanda*, due to the great time depth involved. An alternative explanation for the poor occurrence of the perfect *vs* present *Ablaut* in most IE languages is proposed by Marcantonio. In her opinion, both the perfect *vs* present alternation and the IE *Ablaut* in general may not be an ancient, residual IE phenomenon, but simply an artefact of the method of analysis, as is suggested by the clear lack of 'statistical significance' in the relevant data. Compare also Kazanas' view, according to which only Sanskrit possesses a proper *Ablaut.*

Chapter VI. Paolo Di Giovine, in his chapter: *Verbal inflection from "Proto-Indo-European" to the Indo-European languages: A*

matter of coherence? intends to re-visit that coherent bundle of grammatical features which join together Old Indian, Avestan, Ancient Greek, Gothic, Old Church Slavic, Latin and Hittite, and on which the reconstruction of a PIE verbal system is traditionally based. Against this background, the Author observes, several other grammatical features, several 'independent innovations' can also be identified, whose close scrutiny would show to what extent a different feature "springs up as an independent innovation and to what extent it points to different archetypes" - the latter case in turn would suggest that PIE was not a "compact language, or not even a language at all". In particular, Di Giovine focuses on the still hotly debated issue of the innovative *vs* conservative character of the verbal system of the various IE languages, ranging from the Greco-Aryan verbal system at one end of the spectrum to the Hittite one, at the other. The Author comes to the conclusion that looking at this issue in terms of the usual, simple opposition of: 'archaism *vs* innovation', can be misleading. Similarly, the schematic, binary family tree model is not of much help to properly account for the actual distribution and nature of the morphological (particularly verbal) correlations among the IE languages (as indeed recognized nowadays by many scholars, also in this volume). The type of stratified model of IE put forward in this volume by Drinka and Andersen, according to Di Giovine, does not appear to be satisfactory either. Instead, the model of interpretation that would allow us to 'explain' the morphological correlations in a more realistic way – although by no means in a totally satisfactory way – is the one that depicts IE as a pool of dialectal varieties, a pool of languages that progressively develop and differentiate from one another according to each language internal, specific mechanisms, specific 'drift'. For example, it would be inadequate to consider the Greco-Aryan verbal system purely and simply as an innovation, given that those languages that possess a less complex verbal morphology actually show relics of ancient verbal categories that are no longer productive, or are extinct. This is the case of the Latin *perfectum* (the result of the convergence of aorist and perfect), or the Latin conjunctive of the type: *sīm* and *velim* (formally a continuation of the optative

suffix, which is no longer attested in this language). Thus, a detailed investigation of the history of the single languages /language groups would be required in order to achieve "coherent" and reliable results. Such an investigation would reveal how it is the 'simplification' of the inflectional system that represents innovation, rather than the other way round. This view is held in this volume also by Kazanas, although he reaches this conclusion through different arguments and a different approach to reconstruction.

Chapter VII. Bridget Drinka, in her contribution: *Stratified reconstruction and a new view of the family tree model*, focuses on the debate of whether and how we can sort out genetic *vs* areal correlations, and connected issues of how to interpret the conventional concepts of IE proto-language and family tree. Drinka starts from the assumption that linguistic evidence is an indispensable resource in cultural reconstruction, one which provides more complete information as to the relatedness of particular groups than the archeological record can furnish. The validity of IE as a construct is not contested here, nor is the importance of comparative reconstruction as a valuable tool. What is contested, however, is the over-simplistic characterization of PIE (still maintained by a number of scholars) as a single unified entity, or as an abstraction which serves only to represent the relationship of the daughter languages to one another (not a language which was truly spoken). Drinka presents a critical interpretation of the comparative data, with an eye to constructing a more realistic picture of the development of the IE languages, including stratification of layers and areal contact as important considerations. Specifically, Drinka claims that the conventional genetic model is only endorsed in her analysis as a 'descriptive' model of the ancient relationship of IE languages, but its usefulness as an *explanatory* model is questioned. What the family tree represents is a skeleton, providing a mapping of common ancestry, but no information as to how or why the languages diverged. As it has been amply illustrated since the time of Schmidt (1872), linguistic change occurs through contact, through

the spread of innovation across populations. Hence, some similarities among languages will represent inherited features whilst others will represent shared innovations, acquired through contact (on this topic compare also the chapter by Andersen and Matras). Recent work on languages in contact are claimed (by the Author) to help linguists to sort out the data as archaic or innovative, and to construct a stratified view of PIE accordingly. An updated model which attempts to combine the 'family tree' and the 'wave' models is then proposed. The Author argues that, adopting this new integrated model, it is possible to sort out those morphological similarities (such as iterative, inchoative and intensive affixes) that represent an archaic relationship between the IE languages, from those similarities (such as the development of the imperfect) that represent instead later, shared innovations. The stratification of archaic *vs* innovative structures would allow us to recognize not only that the IE languages are genetically related, but also that some IE languages, like Indo-Iranian and Greek, must have remained in contact longer than others, like Hittite and Germanic. This is testified by the fact that, for example, Indo-Iranian and Greek possess an extremely similar temporal-aspectual system, whilst both Hittite and Germanic have extremely simple verb morphologies. Thus, the shared morphological complexity of Indo-Iranian and Greek would not represent archaism, as traditionally assumed, but rather shared innovation, due to late contacts in the "eastern" areas of IE. In this regard, please compare the chapter by Di Giovine and Kazanas, who support instead the traditional model, although through different corpora of data and different arguments. Compare also the chapter by Schmitt, who argues that (Vedic) Sanskrit is not necessarily the oldest language within IE. Finally, the thesis of late contacts between Greek and Indo-Iranian, according to Drinka, constitutes evidence that the 'Out of India' theory, put forward by Kazanas in this volume, is untenable. However, both Kazanas and Drinka concord in recognizing the validity of IE not only as a linguistic construct, but also as the expression of a real, pre-historical speech community.

Chapter VIII. In his contribution: *The origin and spread of the Indo-Germanic people*, Alexander Häusler states that IE has only a linguistic meaning, and therefore it is methodologically wrong to identify IE or the Indo-Europeans with archaeological cultures, or anthropological types or races. The Author backs his claim through the results of recent archaeological and anthropological research. As a matter of fact, Häusler argues, there is clearly no archaeological or anthropological evidence in favour either of the popular thesis of the existence of a restricted, ancient *Urheimat* or of an *Urkultur* of the Indo-Europeans – see however the chapter by Kazanas, who has a different stand regarding these issues. Similarly, there is no extra-linguistic evidence in favour of the mainstream, migrationist-type of theories, according to which the Indo-Europeans would have dispersed from their *Urheimat* located in the area of the South Russian or Pontic Steppe. On the contrary, at least with regard to the region extending between the North Sea and the Caspian Sea, there is only evidence for a continual, progressive development of the old-established population and culture – undisturbed by migration and conquest – since the Mesolithic / Neolithic times. On the other hand, the Author observes, the so-called 'diffusionist' theses, according to which the old IE population would have diffused peacefully and slowly, due to their attained economical advantages, do not appear to contradict the archaeological evidence. Even so, the (assumed) IE linguistic community is not at all identifiable by the means of archaeological or anthropological research, although it must have existed, since the corresponding linguistic classification has been fairly safely established. This chapter interacts mainly with the chapter by Bryant, Kazanas and Drinka, all dealing (also) with the issue of the IE home land(s) and related migrations, and all holding (more or less) different views.

Chapter IX. Nicholas Kazanas, in his: *Indo-European linguistics and Indo-Aryan indigenism*, argues in favor of the indigenist, 'Out-of -India' hypothesis – as against the Aryan Invasion / Immigration Theory – although he also holds the possibility that the IE home land might have been a very wide area, stretching

from North West India up to the Pontic Steppe. The Author bases his claim on data and argumentations drawn from linguistics (particularly from the lexical, morpho-phonological and morphological structure of Sanskrit), archaeology and genetics, as well as from the system of cultural and ideological values found in the Vedic Hymns. The Indigenist hypothesis supported in this chapter is in contrast with the mainstream doctrine of the European origin of Sanskrit, which is re-proposed and defended in this volume by Drinka and Schmitt. On this issue please compare also the chapter by Bryant, who argues that there is nothing in Indian archaeology "that in any way ...supports a migration of peoples into the Indian subcontinent during even the broadest of time frames associated with the post-PIE dispersions". Kazanas also claims that Sanskrit represents the oldest language within the family, in this respect holding a view which is line with the traditional model, even if, in Kazanas' opinion, the Vedic language and civilization is actually much older than typically assumed within the conventional framework of IE studies: Sanskrit is as old as to even predate the Harappan civilization (but see Schmitt in this volume, who argues instead that there is no sufficient evidence to uphold this claim). Thus, Kazanas believes that the IE theory is well founded, and that an IE language family /speech community did exist in pre-historical times, in this being again in line with the mainstream (and realist) approach to IE. However, Kazanas distances himself from the mainstream doctrine when he argues the following: a) PIE cannot be reconstructed, not only because the widely assumed 'regular and systematic' sound laws have not really been established, but also because reconstructions are, by their very nature, highly conjectural and non-verifiable; b) most of the conventionally established correspondences are actually 'similarities', that is, they do not, strictly speaking, correspond to each others, but are nevertheless 'similar' in sound and meaning. In this respect, Kazanas' overall assessment of reconstructed IE is in line with that of Marcantonio in this volume, although the interpretation of the relevant data and, consequently, the conclusions drawn by the two Authors differ altogether. In fact, Marcantonio interprets the similarities in question mainly as

'chance resemblances', which in turn would be an indication that the IE theory is not, after all, as well founded as widely claimed. In contrast, Kazanas interprets these similarities as 'cognates', because the possibility of chance resemblances or of borrowing is to be excluded given the wide distribution of most of these similarities across Eurasia – but not, significantly, in the Near East. Similarly, according to Kazanas, the well known, extensive similarities of morphological elements and paradigms would be difficult to account for through chance, or borrowing (although there are attested instances of borrowed, wholesale morphological paradigms, as reported in the chapter by Matras).

Chapter X. The main thesis of Angela Marcantonio's chapter: *Evidence that most Indo-European lexical reconstructions are artefacts of the linguistic method of analysis,* is that the methods of historical linguistics, including the comparative method, can be so flexible – by their very nature – that they can be stretched to account for almost any data. This means that the explanatory system runs the risk of becoming dangerously circular, and, therefore, of yielding misleading results – in this case within the field of IE studies. Over the course of about two hundred years of everyday practice of reconstruction within IE the encountered counter-evidence has been typically 'explained away' through all sorts of (often ad-hoc) justifications, so that today one can find hardly any evidence counter to the model. The Author examines this 'explaining away' process critically, by asking the question: have we fitted the IE data to the model, or have we made the model so flexible that it can fit almost any data, including potential counter evidence? Marcantonio analyzes the full comparative corpus of the verbal root reconstructions contained in the most recent IE etymological dictionary (**LIV**), by adopting simple, quantitative methods of analysis. She argues that the great majority of these reconstructed roots lack the required 'statistical significance', the required 'cumulative effect'. As a consequence, the cognates relating to these roots are to be interpreted as 'similarities', rather than 'correspondences'; in turn, these similarities can be interpreted as instances of 'chance

resemblances'. Marcantonio also argues that adding laryngeal segments to the process of reconstruction (whatever the rights and wrongs of the laryngeal theory may be), dangerously increases the explanatory power of the comparative method, in this way further contributing to the flexibility of the overall explanatory model. This chapter interacts mainly with the chapter by Andersen and Kazanas. For example, Kazanas objects to Marcantonio that the scientific method is not applicable to this field of studies – actually, the proper 'scientific' approach would be to ignore rigidity, regularity and uniformity, since the IE linguistic and cultural changes occurred in diverse ways and certainly under no observable laws. Similarly, Andersen believes that scientific proves cannot really be delivered in a field such as this, where scholars cannot observe the objects of investigation directly. Also Drinka and Di Giovine (in this volume) defend the validity of the IE theory, despite their criticism of the one or the other traditional IE tenets. Finally, the reader may want to compare the analysis of the present *vs* perfect *Ablaut* offered in this chapter as against the one proposed by Carruba in his contribution. Both Carruba and Marcantonio point out the scanty and irregular distribution of the vowel alternations across the IE area (with the exception of Old Indian and, partially, Greek). However, Carruba interprets this fact as the fading away through time of a very archaic, IE phenomenon, whilst Marcantonio emphasizes the lack of 'statistical significance' of the phenomenon itself and interprets it as linguistic artefact.

Chapter XI. In: *Defining the limits of grammatical borrowing*, Yaron Matras examines two 'mixed' languages – Romani and Domari – whose populations are known to have come into contact with several other populations and languages in historical times, and it is therefore possible to trace the borrowing that has occurred. These dialects show indeed extensive borrowing, both at the lexical and grammatical level. Close examination of the borrowed grammatical elements shows that some categories are more likely to be borrowed: discourse markers, connectors, focus particles, indefinite expressions, indefinite markers, and

expressions of modality; whilst other categories tend to be borrowed less easily, such as: nominal derivation, plural formation, prepositions, lower numerals, negation and interrogatives. Despite these patterns, no areas or categories of lexicon or of grammar appear to be completely immune from borrowing. Matras presents evidence that case and verbal inflections tend to be borrowed wholesale, if they are borrowed at all. For example, the Greek-derived tense markers in Romani are characterized by a "wholesale import of the entire paradigm, rather than of individual markers, and, in all cases, the imported verb paradigm exists in complementary distribution alongside the inherited inflection paradigm". This evidence illuminates one of the most striking aspects of IE: how does one explain the 'wholesale correlations', spanning entire paradigms, which are observed across some IE languages? For example, textbooks frequently cite tables illustrating the verb 'to carry, to bear', which shows amazing correspondences across Sanskrit, Greek, Latin, Gothic, O. Irish, etc., in most forms of the declension of the present tense, as well as in the lexical forms (compare Sanskrit. *bhárāmi*, Latin *ferō*, Greek φέρω, Gothic *baira*, O Irish **berū* 'I carry', etc.). The verb 'to be' also shows similar – lexical and morphological – correlations across the entire paradigm. There has been an underlying assumption, upon which the IE theory was founded, that these patterns cannot be borrowed and therefore they must be inherited. For example, Nichols (1996a) says that these patterns are "individual-identifying evidence", that is, evidence "obvious to the naked eye of a trained scholar" that the languages are genetically related (this interpretation is supported in this volume by Kazanas). However, there is a minority view, in which Authors question this assumption. For example, Greenberg (2005) and Croft (2005: XIX) remark that these correlations are too good to be the result of inheritance, because of the language change that would have occurred in the millennia involved. Matras' study provides attested, concrete evidence that the borrowing of entire verbal forms – lexical item + paradigm – does indeed occur. This suggests that the familiar examples of verbal paradigms cited in textbooks could equally well be accounted for through borrowing,

and this, in turn, calls into question one of the main, traditional tenets of the IE theory.

Chapter XII. Rüdiger Schmitt, in his: *Iranian archaisms vs. Vedic innovations - and the Indo-Iranian unity*, argues that with regard to the comparative study of the IE languages altogether and to the reconstruction of the IE proto-language it is not sufficient and adequate to refer to the (Old) Indo-Aryan data alone. The wealth and diversity of the textual evidence written in younger Iranian languages as well as the archaism particularly of the Old Avestan language make it only too clear that despite the old age of the earliest Vedic texts the Old Indo-Aryan language is not the only fundament of the IE proto-language, that its data are not with absolute necessity genuinely antique and therefore Old Indo-Aryan (both Vedic and Sanskrit) is not as close to PIE as several people still think. Rather, scholars should take into account the "(proto)-Indo-Iranian linguistic unity", which is indeed "defined by a large number of phonological, morphological, syntactical and phraseological isoglosses found exclusively in both these language families, including the speakers' self-designation as *Aryan*". The Author puts forward a long list of data and argumentations (mainly from the level of phonology), which would show how the Iranian branch often reveals more archaic traits than the Indo-Aryan one. In particular, Schmitt sets himself the task of drawing the list of those Iranian archaisms "that are faced with secondary changes, and are therefore obvious innovations, on the Vedic side". This list, it seems, has never been drawn thus far, despite its obvious relevance for the issues under discussion. Schmitt's findings are at odds with one of the major, traditional tenets of IE studies, that is, that Vedic Sanskrit is the oldest language within the family, and therefore the most important for the process of reconstruction. This tenet is supported in this volume by Kazanas and, to a certain extent, by Di Giovine – the reader might want to compare also Drinka's chapter regarding the status of the Indo-Iranian branch. In particular, Kazanas draws a list of data and features which, in contrast to Schmitt's list, would demonstrate how Vedic is actually more archaic than (Old) Iranian, and the most archaic language within IE. Kazanas also points out that Schmitt's conclusions rely entirely on linguistic data – amenable to a reverse interpretation – and ignore other aspects relevant to the debate, such as evidence drawn from mythology, archaeology and genetics. The claimed archaism of Old Iranian also gives Schmitt the opportunity to consider as correct the mainstream doctrine (re-affirmed in this volume by Drinka and Annamalai & Steever), according to which the Indo-Aryan language must have immigrated into India, rather than having originated there. For an alternative point of view, please compare again the conclusions reached by Kazanas, who, whilst not denying the existence of an Indo-Iranian linguistic unity, argues that this unity must have formed and existed for some time in the Saptasindhu area, whence later on the Iranians left.

6. Conclusion

The reader has seen in this book a variety of views about IE, ranging from the belief that it represents the language of a real pre-historical community; through the thesis that it is only a model to embody linguistic correlations; all the way to statistical evidence that (many) linguistic correlations themselves may be merely an artefact of the method of analysis. In fact, when the various components of the theory are brought together so that they can be seen holistically, it is hard to pin down what the foundations of the theory are actually supposed to be. For example, one of the founding principles of the traditional version of the theory was the assumption that morphological paradigms cannot be borrowed, and therefore it is possible to trace genetic inheritance through them. However, we have seen evidence of wholesale paradigm borrowing, based on studies of languages in contact. In any case, some scholars now hold that morphology is less relevant than other factors – but it is at present unclear whether, or how, these other factors may be verified or falsified. It has been the purpose of this book to bring to the fore these contradictions and open questions associated with the theory. It is for the reader and the linguistic community to decide the way forward.

Endnotes

[1] I have used here the words by H. Andersen, personal communication; 2006.
[2] See the following quote by Blust (1996:137) with regard to the Austronesian languages of island Southeast Asia: "Pandemic irregularity resist a plausible explanation through borrowing, analogy or other mechanisms of secondary change and hence are particularly difficult to reconcile with the Neo-grammarian position that all apparent diachronic rule violations are due to secondary factors which interfere with the perfect regularity of primary change".
[3] According to Ringe et al. (2002), the same would apply to morphology, where, for example, often it is impossible to discover which inflectional markers are ancestral and which may represent innovation.
[4] As pointed out by Hock (1986:633) and Belardi (2002:307), the neo-grammarians worked on sound changes only *ex post facto*; therefore they were unable to observe particular changes in progress; see Kazanas in this volume.
[5] For example, Sihler (1995:157) says: "Much effort has been devoted to trying to discover the conditions under which $*k^w > f$ might be a regular Germanic development, but without success. The probable explanation is that these forms are dialect borrowings from an otherwise un-attested *P*-dialects of Germanic".
[6] Kazanas reports a detailed and exhaustive list of widely diffused cognates, which include, for example, the term for 'blood', 'daughter', 'arm', 'daughter-in-law', 'house', 'metal', 'widow', etc.
[7] Compare for the singular: OIA *-āmi* > Romani *-av*; OIA *-asi* > Romani *-es*; OIA *-ati* > Romani *-el*. For the plural: 1st OIA *-āmas* > possibly Romani *-as*; for the 3rd person OIA *-anti* > Romani *-en*, spreading also to the 2nd person by analogy (Matras 2002: 43).
[8] there are basically two contrasting theories: a) the 'Schwundhypothese', according to which the features and categories present in most /all IE languages but missing in Anatolian (feminine gender, aorist, dual, etc.), have been lost; b) the "Herkunfthypothese", according to which those features and categories absent from Anatolian have been formed only after Anatolian split off from the rest of the family, in which case Anatolian would be particularly archaic. Either of these analyses is void of problems; see Zeilfelder (1994 / 2001: 9-20).
[9] Carruba (2001a:6) states that there is still one knot to untangle: the Indo-Hittite knot, although this should take place in 'a more modern, less polemical manner', than in the twenties or thirties.
[10] Among the various articles contained in this volume see the one by Lehrman, who observes the following (Lehrman 2001:107): "What I find in 2000 is that many scholars have accepted the Indo-Hittite theory in some form – either in name as well as in essence, like Don Ringe (e.g., Ringe 1998b) and Anna Morpurgo Davies (private communication, January 28, 2000, Berkeley, Calif.), or in essence if not in name, as Craig Melchert has done [(2001)]".

[11] Campbell (2003a:11) states that: "Highly recommended though such grammatical evidence is, caution in its interpretation is necessary. There are impressive cases of apparent, idiosyncratic grammatical correspondences which, in fact, have non-genetic explanation (accident or borrowing)".
[12] Compare for, example, the following case endings: Locative *nām(a)ni* in Sanskrit and *nōmine* in Latin; for fuller paradigms and other examples of this type compare Kazanas in this volume (par. 3c.).
[13] Hittite has no aorist; it has one preterit in opposition to the present tense; it has only indicative and imperative, no reduplicated perfect and no separate thematic and a-thematic conjugations. The verbs either follow the *-mi* conjugation or the *-ḫi* conjugation, although there are claimed to be a few, good matches to stems, suffixes and even whole paradigms with other IE languages.
[14] It could be argued that the thesis of the 'loss' of rich morphology is a typically 'circular' explanation, even if (one could counter-argue) this explanation appears to be supported by the fact that the modern IE languages have all manifestly lost the original inflected structure, as shown by the modern Indo-Iranian languages, the Romance languages as well as the case of Romani and Domari, for which see again the chapter by Matras in this volume.
[15] The issue of the presence *vs* absence of a rich (mainly inflectional) morphology has long been known; see for example Fortson (2004:224-5); (Meillet 1908); Birwé (1955); Meid (1975); (Neu 1976); Polomé (1981 & 1985); Drinka (1993/95, 1995, 2001 & 2003); etc.
[16] There appears to be evidence, from archaeology and palaeo-anthropology, of pre-historical, as well as, of course, historical contacts between Asia Minor and the Indus Valley. See Sedlar (1980); Tucci (1982); Rawlinson (1916 &1975) and Basham (ed. 1975).
[17] Compare also the isoglosses that separate Hittite and Italic from Sanskrit and Greek (such as augment in past tense and prohibitive Skt. *mā* and Gk. *mē*), or the morphological isogloss of the locative plural ending *-su* / *-si*, which encompasses Sanskrit, Greek and Slavic. On the other hand, palatalization separates Sanskrit from Hittite, Greek and Italic. Notice also that palatalization is not at all a clear-cut isogloss, especially in Baltic and Slavic, for which see Miller (1976).
[18] This problem emerges again in the various IE 'philogenies' attained through 'cladistic' techniques. See Gray & Atkinson (2003); Ringe, Warnow & Taylor (2002); Garrett (2006), Forster & Renfrew (2006), and again Drinka in this volume.
[19] Between October and December 2005 the Faculty of Classics of the University of Cambridge hosted a series of seminars on issues in IE linguistics, organized by A. Marcantonio.
[20] As an example of a well known noun belonging to this class (masculine), compare: Gk. λύκος, Lat. *lupus*, Goth. *wulfs*, OCS *vlĭkŭ*, Lith. *vilkas* 'wolf'; etc.

[21] See Mayrhofer (1973:38-9); Szemerényi (1996:183-7); Sihler (1995: 257-265); Meier-Brügger et al. (2003:198-9); Rix (1992:135 ff.); Fortson (2004:113-7); Clackson (2007: 9-10), etc.

[22] This ending is ambiguous as to its grade.

[23] Hittite, for example, has no nominal declension corresponding to the feminine stems in *-*e*h_2 or **i*-h_2 (as these are currently reconstructed).

[24] The same could be said about the many difficulties and mismatches encountered at the phonological level between Anatolian and the other IE languages (as clearly pointed out by Melchert (1994); Kimball (1999) and Hawkins & Morpurgo Davies [1998]). Compare also Zeilfelder's view (1994/2001:3 ff.), according to which the phonology (and lexicon) of the Anatolian languages is not of much help for the purpose of verifying either the thesis of the archaic or that of the innovative nature of the Anatolian branch.

[25] Justus's article (2002:121-123) offers a summary of "contemporary views" on the problem of 'gender' in Hittite and related attempts to solutions.

[26] Famous heteroclitic *r* /*n*-stem, neuter nouns include: 1) the term for 'water': **wod-r/n-*, with alternating stem **wed-n-*, preserved almost exactly in Hitt. *wātar* → Gen. *witenaš*; 2) the term for 'fire': **peh*$_2$*-wr̥*, continued most faithfully in Hitt. *paḫḫur* → Gen. *paḫḫuenaš*; 3) the term for 'blood': *ēšḫar* → Gen. *ešḫanaš*. Compare also: 4) Lat. *femur* ~ *femin-is* 'thigh'; 5) *iecur* ~ *iecineris* (from earlier *iecor* ~ **iecinis*), the latter corresponding to Skt. *yakn̥* ~ *yakanas* 'liver'. Contrary to what happens in Anatolian, the original mode of inflection is scantily attested elsewhere, in which case 'conflation of paradigms' is typically assumed (see Szemerényi (1996:173); Fortson (2004:110-2) & Clackson [2007: 94-5]).

[27] Compare Fortson (2004:227): "Before the decipherment of Mycenaean, it had been assumed that the labiovelars were still intact in Proto-Greek [...] even though none of the dialects then known actually preserved them as labiovelars. This assumption had been made because it was the easiest way to explain certain consonant alternations within the dialects [...]. Mycenaean proved this hypothesis correct: it still preserved the labiovelars intact, or at least preserved them systematically as something different from labials, dentals, or plain velars".

[28] Since we do not have a segment-by-segment spelling system for Mycenaean, we cannot directly prove that the segment at the beginning of those syllables (where we would reconstruct a labio-velar) is indeed a labio-velar. All one can safely state is that the symbol in question is not the one used in syllables which have /p/, /t/ or /k/ in other Greek dialects; therefore, the initial consonant must be something else.

[29] Hock (1993:221ff.) observes the following: "Preliminary results [of an experiment in which English, Finnish and Hindi are compared] suggest that enlarging the data base does not improve the reliability of the method. In fact, if there is any change at all it may consist in a slight increase of false friends". On the other hand, Hock (*ibidem*) also recognizes that working through a comparison of only a small number of languages (three in the specific case), may be equally misleading.

[30] Personal communication; 2007.

[31] The claim that prothetic vowels can be explained through laryngeals in Greek and Armenian has been contested by Wyatt (1972) and Beekes (1969:18-98).

[32] Compare also the volume edited by Vennemann (1989), where the 'new sound of Indo-European' appears to have had triggered some reshaping in the definition of the laryngeal theory. Furthermore, it is now recognized that Hittite does not really lend much support to the laryngeal theory, since the evidence available from this language is, admittedly, 'fragmentary', 'contradictory', 'disappointing' (see Lehmann (1952: 25 ff. & 1993:125); Szemerényi (1996: 121-126 & 137-8); Mayrhofer (1986:123ff), Clackson (2007:58 ff); etc.). The absence of satisfactory evidence for laryngeals from the Anatolian languages seems to have been compensated, in recent years, through data drawn from Greek, which is now claimed to show an unparalleled 'triple reflex', that is, a distinct outcome for each of h_1, h_2 and h_3, when they occur within a few specific contexts (see Clackson 2007-58-60). This analysis appears to be gaining acceptance, even though laryngeals nowhere survive as consonants in Greek.

[33] Sihler's reconstruction is much leaner than others because it works with only "a SINGLE pattern of alternations, **e* ~ **o* ~ *Ø*", together with "three new consonants of obscure phonetics", the laryngeals *H_1, *H_2 and *H_3. Note however that Sihler does not appear to consider this extra number of rules as a 'disadvantage', that has a negative impact on the 'economicity' of the system. This is, in fact, an interpretation of the Author (A. M.) of the present *Introduction*.

[34] Similar remarks can be found also in Collinge (1985) and Gusmani (1979).

[35] For an up-to-date state-of-the-art report on the laryngeal theory & essential bibliography see: Polomé (1987); Mayrhofer (2004:17-39); Meier-Brügger et al. (2003:106 ff.); Di Giovine (2005) & Clakson (2007); see also Sturtevant (1942).

[36] Last, but not least, one should take into consideration also those further limitations of the comparative method as pointed out in recent years by scholars such as Dixon (1997), Aikhenvald & Dixon (2001) and Nichols (1992, 1993/1995, 1996b, 1997 & 1998); see also Andersen (2006). These are: a) the inability of the method to get at a more remote linguistic prehistory than the generally assumed 8000/6000 years); b) its inability to account for the actual distribution of the degree of linguistic 'diversity' *vs* linguistic 'uniformity' found in the world. These issues are not dealt with in this volume.

[37] For example, Meid (1975), in proposing his 'Space-Time' model, claims that the spoken IE parent language did exist, and that there must have been a corresponding very small, tight-knit community which slowly expanded over a vast area, and through a long stretch of time. On the contrary, Clackson

(2007:16) states that PIE "is not a real language" and that: "Reconstructed IE is a construct which does not have any existence at a particular time and place".

[38] Among the criticisms raised in the literature against the methods of palaeo-linguistics see Renfrew (1987:77-86, 103-104, 109-110) and Marcantonio (2002/5); compare however Polomé (1990) for a rather positive evaluation.

[39] Please, note that although R. Schmitt contributes with a chapter in this volume, his contribution by no account deals with the issue of how real(istic) reconstructions may be.

[40] Hock (1996:38) states that: "given the multiply different possible explanations, no explanation can be considered certain enough to be used as evidence for or against prehistoric Dravidian/Indo-Aryan contacts".

[41] Hock (1993) observes that within well established families too, such as IE, a significant number of chance resemblances may occur, or, conversely, a number of cognates are very hard or even impossible to detect because of the changes they have undergone though time.

[42] The classification of Sanskrit is not the only known instance of linguistic classification which has been influenced by historical /political events. There is a clear, well documented example of this process: Hungarian, for which see Marcantonio 2002.

[43] Clackson (2007: 3) says: "The operation of the comparative method does not guarantee a language's place in the family; only the initial recognition that two or more languages are related can do that [...]. When does a linguist decide that there is enough material to relate a language to the IE family? There is no absolute set of criteria beyond the general rule that the evidence must convince both the individual linguist and the majority of the scholar community".

References

Aikhenvald, A. Y. & Dixon R. M. W. (eds)
2001 *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*. Oxford University Press.
Aikhenvald, A. Y. & Dixon R. M. W.
2001 Introduction. In A. Y. Aikhenvald & R. M. W Dixon (eds). 1-26.
Aiyar, R. S.
1975 *Dravidian Theories*. Madras: Madras Law Journal Office.
Andersen, H.
1988. Centre and periphery: adoption, diffusion and spread. In J. Fisiak (ed.), *Historical Dialectology*. Berlin: Mouton de Gruyter. 39-83.

1996 *Reconstructing Pre-historical Dialects: Initial Vowel in Slavic and Baltic*. Trends in Linguistics. Studies and Monographs, 91. Berlin /New York: Mouton de Gruyter.
1998 Slavic. In A. Giacalone & P. Ramat (eds). 415-54.
2006 Review of: *Examining the farming / language dispersal hypothesis*, ed. by Peter Bellwood and Colin Renfrew (McDonald Institute Monographs). Cambridge: McDonald Institute for Archaeological Research. 2002. *Folia Linguistica Historica* 27:201-207.
Andersen, H. (ed.)
1993/95 *Historical linguistics 1993. Selected Papers form the 11th International Conference on Historical Linguistics*. Current Issues in Linguistic Theory, 124. Amsterdam /Philadelphia: Benjamins.
2003 *Language in Contacts in Pre-history: Studies in S*tratigraphy. Current Issues in Linguistic Theory, 239. Amsterdam /Philadelphia: Benjamins.
Andersen, H. & Koerner, K. (eds)
1987/1990 *Historical Linguistics 1987. Papers from the 8th International Conference on Historical Linguistics*. Current Issues in Linguistic Theory, 66. Amsterdam /Philadelphia: Benjamins.
Annamalai, E.
1985 *Dynamics of Verbal Extension in Tamil*. Trivandrum, Kerala: Dravidian Linguistic association of India.
2001 *Managing Multilingualism in India: Political and Linguistic Manifestations*. London / New Delhi: Sage Publications.
Austin, P.
1989 Switch-reference in Australia. *Language* 57: 309-334.
Austin, P. (ed.)
1990 *Languages and History. Essays in Honour of Luise A. Hercus*. Canberra: Australian National University. Pacific Linguistics. Series C; 116.
Basham, A. L. (ed.)
1975 *A Cultural History of India*. Oxford: Clarendon Press.
Beekes, R. S. P.
1969 *The development of the Proto-Indo-European Laryngeals in Greek*. Paris / The Hague: Mouton.
1989 The nature of the proto-Indo-European laryngeals. In T. Vennemann (ed.). 23-33.
1995 *Comparative Indo-European Linguistics. An Introduction*. Amsterdam / Philadelphia: Benjamins.

Belardi, W.
2002 *L'Etimologia nella Storia della Cultura Occidentale*, I. Roma: Il Calamo.
Benveniste, E.
1935 *Origines de la Formation des Noms en Indo-Européen.* Paris: Librairie Adrien-Maisonneuve.
Birnbaum. H.
1977 Linguistic reconstruction: Its potential and limitations. *Journal of Indo-European Studies*, 20-51.
Birwé, R.
1955 *Griechisch-arische Sprachbeziehungen im Verbalsystem.* Waldorf-Hessen: Verlag für Orientkunde.
Blake, B. J. & Burridge, K. (eds)
2001 *Historical Linguistics 2001. Selected Papers from the 15th International Conference on Historical Linguistics.* Current Issues in Linguistic Theory, 237. Amsterdam /Philadelphia: Benjamins.
Blenvins, J.
2004 *Evolutionary Phonology. The Emergence of sound patterns.* Cambridge University Press.
Blust, R.
1996 The Neo-grammarian hypothesis and pandemic irregularity. In R. Ross & R. Durie (eds). 135-158.
Bonfante, G.
1931 I dialetti indoeuropei. *Annali dell'Istituto Orientale di Napoli* 4: 69-185.
1945 On reconstruction and linguistic method. *Word* 1: 83-94, 132-61.
Boysan-Dietrich, N., Marazzi, M. & Nowicki, H. (eds)
1983 *Sammlung hieroglyphischer Siegel, I.* Würzburg: Königshausen & Neumann.
1985 *Supplementum zu Sammlung hieroglyphischer Siegel, I.* Würzburg: F. Schöning.
Brady, R & Marcantonio, A.
2003 Evidence that most Indo-European lexical reconstructions are artefacts of the linguistic method of analysis. In Hajičová, E. et al. (eds), *Proceedings of the 17th International Congress of Linguists.* Prague: Matfyzpress, MFF UK. CD-ROM ISBN: 80-86732-21-5.
Bryant, E. F.
2001 *The Quest for the Origins of the Vedic Culture. The Indo-Aryan Migration debate.* Oxford University Press.

Bryant, E. F. & Patton L. L. (eds)
2005 *The Indo-Aryan Controversy: Evidence and Inference in Indian History.* London /New York: Routledge.
Burrow, T.
1946 Loan words in Sanskrit. *Transactions of the Philological Society*` I:1-30.
1973 *The Sanskrit Language.* London: Faber & Faber [revised edition].
1975 The early Aryans. In A. L. Basham (ed.), *A Cultural History of India.* Oxford: Clarendon Press. 20-9.
Callaghan, C.
1980 An 'Indo-European' type paradigm in Proto Eastern Miwok. In K. Klar, M. Langdon and S. Silver (eds), *American Indian and Indo-European studies: Papers in Honour of Madison S. Beeler.* The Hague: Mouton de Gruyter. 331-8.
Campanile, E.
1998 The Indo-Europeans: origins and culture. In A. Giacalone & P. Ramat (eds). 1-24.
Campbell, L.
1995 The Quechumaran hypothesis and lessons for distant genetic comparisons. *Diachronica* 12/2: 157-200.
1998 *Historical Linguistics: an Introduction.* Edinburgh University Press.
2001 Beyond the comparative method? In B. J. Blake & K. Burridge (eds). 33-57.
2003a How to show languages are related: methods for distant genetic relationship. In B. D. Joseph & R. D. Janda (eds), *The Handbook of Historical Linguistics.* Oxford University Press.
2003b Historical linguistics: the state of the art. In Hajičová, E. et al. (eds), *Proceedings of the 17th International Congress of Linguists.* Prague: Matfyzpress, MFF UK. CD-ROM ISBN: 80-86732-21-5.
Carruba, O.
1997 *Analecta Linguistica Anatolica.* Pavia: Juculano Editore.
2000 Indo-European **sem/sm* in the pronouns: 'singulative' plurals. *Journal of Indo-European Studies* 28:341-357.
2001a. Premessa. In O. Carruba & W. Meid (eds). 1-7.
2001b Genere e classe in anatolico: La 'mozione in *-i*' e il 'caso in *-sa/-za*'. In O. Carruba & W. Meid (eds). 29-42.
Carruba , O & Meid, W. (eds)
2001 *Anatolisch und Indogermanisch.* Akten des Kolloquiums der Indogermanischen Gesellschaft, Pavia, 22.-25. September 1998. Innsbruck.

Chapman, J. & Hamerow, H. (eds)
1997 *Migrations and Invasions in Archaeological Explanation.* Oxford: BAR International Series 664.

Clackson, J.
2007 *Indo-European linguistics. An Introduction.* Cambridge University Press.

Collinge, N. E.
1985 *The Laws of Indo-European.* Amsterdam: Benjamins.

Comrie , B.
1993 Typology and reconstruction. In C. Jones (ed.). 74-97.
1998 The Indo-European linguistic family: genetic and typological perspectives. In A. Giacalone & P. Ramat (eds). 74-97.

Croft, W.
2005 Introduction. In: Greenberg, J. H. *Genetic Linguistics: Essays on Theory and Method.* Edited by W. Croft. Oxford University Press. 2-55.

Dench, A.
2001 Descent and diffusion: the complexity of the Pilbara situation. In A. Y. Aikhenvald & R.M.W. Dixon (eds). 105-133.

De Saussure, F.
1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes.* Leipzig: Teubner.

Diakonoff, I. M.
1990. Language contacts in the Caucasus and the Near East. In Markey, T. L. & Greppin, A. C. (eds), *When the Worlds Collide. The Indo-Europeans and the Pre-Indo-Europeans.* Ann Arbor, MI: Karoma. 53-65.

Di Giovine, P.
1995 *Studio sul Perfetto Indo-Europeo*, II. Roma: Il Calamo.
2004 Il perfetto indoeuropeo tra endomorfismo ed esomorfismo. In J. H. W. Penney (ed.), *Indo-European Perspectives. Studies in Honour of Anna Morpurgo Davies.* 3-17.
2005 Le laringali indo-europee: un fantasma della ricostruzione? In R. Bombi, G. Cifoletti, F. Fusco, L. Innocente & V. Orioles (eds), *Studi in onore di Roberto Gusmani.* Alessandria: Edizioni dell'Orso. 4-30.

Dixon, R. M. W.
1997 *The Rise and fall of Languages.* Cambridge University Press.
2002 *Australian languages: Their nature and development.* Cambridge University Press.

Drews, R. (ed.)
2001 *Greater Anatolia and the Indo-Hittite Language Family. Journal of Indo-European Studies.* Monograph Series n. 38.

Drinka, B.
1993 /95 Areal linguistics in prehistory: evidence form Indo-European aspect. In H. Andersen (ed.). 143-158.
1995 *The Sigmatic Aorist in Indo-European: Evidence for the Space-Time Hypothesis. Journal of Indo-European Studies.* Monograph series n. 13.
2001 The formation of periphrastic perfects and passives in Europe. In B. J. Blake & K. Burridge (eds). 103-128.
2003 The development of the perfect in Indo-European: Stratigraphic evidence of prehistoric areal influence. In H. Andersen (ed.). 77–105.

Emeneau, M.
1956 India as a linguistic area. *Language* 32: 3-16.
1971 Dravidian and Indo-Aryan: the Indian linguistic area. In A. F. Sjoberg (ed.), *Symposium on Dravidian Civilization.* Austin (Texas): Jenkins. 33-68.
1980 *Language and Linguistic Area. Essays Selected by A. S. Dil.* Stanford (CA): University Press.

Forster, P. & Renfrew, C. (eds)
2006 *Phylogenetic Methods and the Prehistory of Languages.* Cambridge: McDonald Institute for Archaeological Research.

Fortson, B. W.
2004 *Indo-European Language and Culture: An Introduction.* Oxford: Blackwell.

Fox, A.
1995 *Linguistic Reconstruction. An Introduction to the Theory and Method.* Oxford University Press.

Friedrich, J.
1924 Die bisherigen Ergebnisse der hethitischen Sprachforschung. In *Stand und Aufgaben der Sprachwissenschaft.* Fs Streitberg. Heidelberg. 204-218.

Friedrich, P.
1975 *Proto-Indo-European Syntax. The Order of Meaningful Elements. Journal of Indo-European Studies.* Memoir 1. Butte (Montana): College of Mineral Science.

Garrett, A.
2006 Convergence in the formation of Indo-European Subgroups: Phylogeny and Chronology. In Forster, P. & Renfrew, C. (eds). 139-5.

Giacalone, A. & Ramat, P. (eds)
1998 *The Indo-European Languages.* London /New York: Routledge.

Gimbutas, M.
1970 Proto-Indo-European culture: The Kurgan culture during the fifth, fourth, and third millennia BC. In G. Cardona, H. Hoenigswald & A. Senn (eds), *Indo-European and the Indo-Europeans: Papers presented at the Third Indo-European Conference at the University of Pennsylvania.* Philadelphia: University of Pennsylvania Press. 155-197.

Grace, G.
1990 The "aberrant" versus "exemplary" Melanesian languages. In P. Baldi (ed.), *Linguistic Change and Reconstruction Methodology.* Berlin: Mouton de Gruyter. 155-173.
1996 Regularity of change in what? In M. Ross & M. Durie, (eds). 157-179.

Grassman, H.
1872 *Wörterbuch zum Rig-Veda.* Wiesbaden: Harrassowitz.

Gray, R. D. & Atkinson, Q.D.
2003 Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* 426: 453-9.

Greenberg, J. H.
1987 *Language in the Americas.* Stanford University Press.
1991 Some problems of Indo-European in historical perspective. In S. M. Lamb & E. D. Mitchell (eds), *Sprung from Some Common Source: Investigation into the Prehistory of Languages.* Stanford University Press. 127-40.
2000 *Indo-European and its Closest relatives: The Eurasiatic Language Family.* Vol.I: Grammar. Stanford University Press.
2001 *Indo-European and its Closest relatives: The Eurasiatic Language Family.* Vol. II: Lexicon. Stanford University Press.
2005 *Genetic Linguistics: Essays on Theory and Method.* Edited by W. Croft. Oxford University Press.

Gusmani, R.
1979 Ittito, teoria laringalistica e ricostruzione. In E. Neu & W. Meid (eds). 63-72.

Harrison, S. P.
2003 On the limits of the comparative method. In B. D. Joseph & R. D. Janda (eds). 213-243.

Häusler, A.
1999 Nomadenhypothese und Ursprung der Indogermanen. In E. C. Polomé (ed.), *Miscellanea Indo-Europea. Journal of Indo-European Studies,* Monograph Series n. 33. 117-170.
2002 Ursprung und Ausbreitung der Indogermanen: Alternative Erklärungsmodelle. *Indogermanische Forschungen* 107: 47-75.

2003a Bemerkungen zu einigen Ansichten über den Ursprung der Indogermanen. *General Linguisitcs* 40. *Indo-European Language and Culture: Essays in Memory of Edgar C. Polomé.* 131-47.
2003b *Nomaden, Indogermanen, Invasionen. Zur Entstehung eines Mythos.* Halle: Orientwissenschaftliches Zentrum der Martin-Luther-Universität Halle-Wittenberg.
2003c Urkultur der Indogermanen und Bestattungsriten. In A. Bammesberg & T. Vennemann (eds), *Languages in Prehistoric Europe.* Heidelberg: Winter. 49-83.
2004 Problems of the origins of the Indo-Europeans. *Acts of the XIV UISPP Congress. University of Liège, Sptember 2001.* Liège: BAR International Series 1302. 79-84.

Hawkins, J. D. & Morpurgo Davies, A.
1998 Of donkeys, mules and Tarkondemos. In J. H. Jasanoff, C. H. Melchert & L. Oliver (eds), *Mír Curad. Studies in Honour of Calvert Watkins.* Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 243-60.

HED = Puhvel, J.
1984-2001 *Hittite Etymological Dictionary.* Berlin /New York: Mouton.

Hock, H. H.
1975 Substratum influence on (Rig-Vedic) Sanskrit? *Studies in the Linguistic Science,* 5/2: 76-125.
1986 *Principles of Historical Linguistics.* Berlin /New York: Mouton de Gruyter.
1993 SWALLOW TALES: Chance and the "world etymology" MALIQ'A 'swallow, throat'. In K. Beals et al. (eds), *Papers from the 29th Regional meeting of the Chicago Linguistic Society.* 215-238.
1996 Pre-rigvedic convergence of Indo-Aryan and Dravidian? A survey of the issues and controversies. In J. E. M. Houben (ed.), *Ideology and Status of Sanskrit.* Leiden: Brill. 17-58.
1999 'Out of India? The linguistic evidence. *Aryan and Non-Aryan in South Asia.* Harvard Oriental Series Opera Minora 3. 1-18.

Jasanoff, J. H.
2003 *Hittite and the Indo-European Verb.* Oxford University Press.

Jones, C. (ed.)
1993 *Historical Linguistics: Problems and Perspectives.* London / New York: Longman.

Jones-Bley, K. & Zdanovich, D.G.
2002 *Complex Societies of Central Eurasia from the 3rd to the 1st Millennium BC. Regional Specifics in Light of Global Models.*

Journal of Indo-European Studies, Monographs Series n. 45 & 46.

Joseph, B. D. & Janda, R. D. (eds)
2003 *The Handbook of Historical Linguistics*. Oxford: Blackwell.

Justus, C. F.
2002 Hittite and Indo-European gender. In M. R. V. Southern (ed.), *Indo-European Perspectives. Journal of Indo-European Studies*. Monograph Series n. 43. 121-150.

Kazanas, N.
2001 Indo-European deities and the Rigveda. *Journal of Indo-European Studies* 29: 257-93.
2002 Indigenous Indo-Aryans and the Rigveda. *Journal of Indo-European Studies* 30: 270-334.
2003 Final Reply. *Journal of Indo-European Studies* 31:189-196.
2004 Sanskrit and proto-Indo-European. *Indian Linguistics* 65: 75-100.

Kimball, S. E.
1999 *Hittite Historical Phonology*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Koch, H.
1996 Reconstruction in morphology. In M Ross & M Durie (eds). 218-63.
2001 Morphological reconstruction as an etymological method. In B. J. Blake & K. Burridge (eds). 271-291.

Koerner, K.
1989 Comments on reconstruction in historical linguistics. In T. Vennemann (ed.). 3-17.

Kuiper, F. B.
1955 Rigvedic loan words. In O. Spies (ed.), *Studia Indologica: Festschrift für Willibald Kirfel zur Vollendung seines 70. Lebensjahres*. Bonn: Orientalistisches Seminar der Universität. 137-185.
1991 *Aryans in the Rigveda*. Leiden Studies in Indo-European, 1. Amsterdam /Atlanta: Rodopi.
1992 Rigvedic loanwords. *International Journal of Dravidian Linguistics*. 21: 1-49.

Kuzmina, E. E.
2003 The genesis of the Indo-Aryans in the light of data of historical tradition and archaeology. In K. Jones-Bley et al. (eds), *Proceedings of the 15th Annual UCLA Indo-European Conference*. JIES, Monograph Series 49.

Lass, R.
1993 How real(ist) are reconstructions? In C. Jones (ed.). 156-189.

1997 *Historical linguistics and Language Change*. Cambridge: Cambridge University Press.

Lazzeroni, R.
1998 Sanskrit. In A. Giacalone & P. Ramat (eds). 98-124.

Lehmann W. P.
1952 *Proto-Indo-European Phonology*. Austin: University of Texas Press.
1993 *Theoretical Bases of Indo-European Linguistics*. London: Routledge.
2001 Reconstructing Proto-Indo-Hittite. In R. Drews (ed.). 106-130.

Levin, S.
1995 *Semitic and Indo-European. The Principal Etymologies*. Amsterdam Studies in the Theory and History of Linguistic Science. Amsterdam / Philadelphia: Benjamins.

Lindeman, F. O.
1982 *The triple Representation of Schwa in Greek and some Related Problems of Indo-European Phonology*. Oslo-Bergen-Tromsö: Universitetsfolaget.

LIV = Rix, H.
1998 *Lexikon der Indogermanischen Verben. Die Wurzeln und und ihre Primärstammbildungen*. Wiesbaden: L. Reichert.

Luraghi, L.
1998 The Anatolian languages. In A. Giacalone & P. Ramat (eds). 168-196.

Mallory, J. P.
1989 *In Search of the Indo-Europeans. Language, Archaeology and Myth*. London: Thumes & Hudson.
2001 Gli Indeuropei ed i popoli delle steppe: il modello della sostituzione delle lingue. *Proceeding of the Congress: Le Radici Prime dell'Europa. Gli Intrecci Genetici, Linguistici, Storici*. Milano: Paravia – Mondadori. 138-164.

Marazzi, M. et al.
1990 *Geroglifico Anatolico: Problemi di Analisi e Prospettive di Ricerca*. Roma: Biblioteca di Ricerche Linguistiche e Filologiche 24.

Marcantonio, A.
2002/05 Linguistic palaeontology: science or fiction? A case study in Uralic. In B. Winsa (ed.), Finno-Ugric Peoples in the Nordic Countries. *Proceeding of the 'V ROOTs' Conference: the Roots of peoples and Languages of Northern Eurasia*. Stockholm: Academia Tornedaliensis. 77-94.
2002 *The Uralic Language Family: Facts, Myths and Statistics*. Oxford/Boston: Blackwell.

2003/05 Evidence: the missing concept in comparative studies. A preliminary comparison of Uralic and Indo-European. In M. M. J. Fernandez-Vest (ed.), *Les Langues Ouraliennes Aujourd'hui: Approche Linguistique et Cognitive / The Uralic Languages Today: a Linguistic and Cognitive Approach.* Bibliothèque de l'École des Hautes Études: Sciences Historiques et Philologiques 340. Paris: Librairie Honoré Champion. 117-132.

Martinet, A.
1953 Non-apophonic *o* in IE. *Word* 9: 253-67.

Masica, C. P.
1976 *Defining a Linguistic Area: South Asia.* Chicago University Press.
1991 *The Indo-Aryan Languages.* Cambridge University Press.

Matras, Y.
2002 *Romani: A Linguistic Introduction.* Cambridge University Press.

Matras, Y & Bakker P. (eds)
2003 *The Mixed Languages Debate: Theoretical and Empirical Advances.* Berlin / New York: Mouton de Gruyter.

Mayrhofer, M.
1973 *A Sanskrit Grammar.* Alabama Linguistics and Philological Series, n. 20. The University of Alabama Press.
1986 *Indogermanische Grammatik I-2. Lautlehre. (Segmentale Phonologie des Indogermanischen).* Heidelberg: Winter.
2004 *Die Hauptprobleme der indogermanischen Lautlehre seit Bechtel*: Wien: Der Österreichischen Akademie der Wissenschaften.

McMahon, A.
2005 Introduction. *Transactions of the Philological Society* 103 /2: 113-119.

McMahon, A. & McMahon, R.
2003 Finding families: quantitative methods in language classification. *Transactions of the Philological Society* 101/1: 7-57.
2005 *Language Classification by Numbers.* Oxford University Press.

Meid, W.
1975 Probleme der räumlichen und zeitlichen Gliederung des Indogermanischen. In H. Rix (ed.), *Flexion und Wortbildung.* Wiesbaden: Reichert. 204-19.

Meier-Brügger, M., Fritz, M. & Mayrhofer, M.
2003 *Indo-European Linguistics.* Berlin / New York: Mouton de Gruyter.

Meillet, A.
1908 *Les Dialectes Indo-Européens.* Paris: Honoré Champion.
1934 *Introduction à l'Étude Comparative des Langues Indo-Européennes.* Paris: Librairie Hachette.

Melchert, H. C.
1992 Relative chronology and Anatolian: the vowel system. In R. Beekes, A. Lubotsky & J. Weitenberg (eds), *Akten der VIII. Fachtagung der Indogermanischen Gesellschaft.* Wiesbaden: Reichert. 42-35.
1994 *Anatolian Historical Phonology.* Amsterdam & Atlanta: GA: Rodopi.
2001 Critical response to the last four papers. In R. Drews (ed.). 229-36.

Miller, D.G.
1976 Pure velars and palatals in Indo-European: a rejoinder to Magnusson. *Linguistics* 178: 46-64.

Misra, S. S.
2005 The date of the Rigveda and the Aryan migration. In E. F. Bryant & L. L. Patton (eds). 181-233.

Morpurgo Davies, A.
1978 Analogy, segmentation and early Neogrammarians. *Transactions of the Philological Society. Commemorative volume: The Neo-grammarians.* 36-64.
1998 Nineteenth-Century Linguistics. In Lepschy, G. (ed.), *History of Linguistics* IV. London: Longman.

Neu, E.
1976 Zur Rekonstruktion des indogermanischen Verbalsystems. In A. Morpurgo Davies & W. Meid (eds), *Studies in Greek, Italic and Indo-European Linguistics Offered to Leonard R. Palmer.* Innsbrucker Beiträge zur Sprachwissenschaft.

Neu, E. & Meid, W. (eds)
1979 *Hethitisch und Indogermanisch. Vergleichende Studien zur historischen Grammatik und zur dialektgeographischen Stellung der indogermanischen Sprachgruppe Altkleinasiens.* Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck

Nichols, J.
1992 *Linguistic Diversity in Space and Time.* Chicago University Press.
1993/95 Diachronically stable structural features. In H. Andersen (ed.). 337-56.
1996a The comparative method as heuristic. In Ross, M. & Durie, M. (eds). 39-71.

1996b The geography of language origins. *Berkley Linguistic Society* 22: 262-78.

1997 Modelling ancient population structures and movement in linguistics. *Annual Review of Anthropology* 26: 359-84.

1998 The Eurasian spread zone and the Indo-European dispersal. In R. Blench & M. Sprigs (eds), *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*. London: Routledge. 220-66.

Oettinger, N.

1987 Bemerkungen zur anatolischen *i*-Motion und Genusfrage. *Zeitschrift für Vergleichende Sprachforschung*, 100: 35-43.

Pisani, V.

1933 *Studi sulla Preistoria delle Lingue Indoeuropee*. Roma: Atti della Reale Accademia Nazionale dei Lincei. Memorie. Serie VI, Vol. IV, Fasc. VI.

1940 *Geolinguistica e Indoeuropeo*. Roma: Atti della Reale Accademia Nazionale dei Lincei. Memorie. Serie VII, Vol. IX, Fasc. II.

1974 *Indogermanisch und Europa* (*Ars Grammatica*, Band 2). München: W. Fink Verlag.

Polomé, E. C.

1981 Indo-European verb morphology. An outline of some recent views, with special regard to Old Indic. *The International Journal of Dravidian Languages* 9 /1: 158-69.

1985 How archaic is Old Indic? In U. Pieper & G. Sticker (eds), *Studia Linguistica Diachronica et Synchronica*. Berlin /New York: Mouton De Gruyter.

1987 Recent Developments in the *Laryngeal* Theory. JIES 15:159-67.

1990 Linguistic palaeontology: Migration theory, prehistory, and archaeology correlated with linguistic data. In E. C. Polomé (ed.). 137-59.

Polomé, E. C. (ed.)

1990 *Research Guide on Language Change*. Trends in Linguistics. Studies and Monographs, 48. Berlin/New York: Mouton de Gruyter

Pulgram, E.

1959 Proto-Indo-European reality and reconstruction. *Language* 35: 421-26.

1961 The nature and use of proto-languages. *Lingua* 10: 18-37.

Rawlinson, H. G.

1916 *Intercourse between India and the Western World: from the Earliest times to the fall of Rome*. Cambridge University Press.

1975 Early contacts between India and Europe. In A. L. Basham (ed.). 426-438.

Renfrew, C.

1987 *Archaeology and Language. The puzzle of Indo-European Origins*. London: J. Cape.

1990 Archaeology and linguistics: some preliminary issues. In Markey, T. L. & Greppin, A. C. (eds), *When the Worlds Collide. The Indo-Europeans and the Pre-Indo-Europeans*. Ann Arbor, MI: Karoma. 15-25.

Ringe, D.

1992 On Calculating the Factor of Chance in Language Comparison. *Transactions of the American Philosophical Society* 82: 1-110. Philadelphia.

1993 A reply to Professor Greenberg. *Proceedings of the American Philosophical Society* 137/1: 91-109.

1995 "Nostratic" and the factor of chance. *Diachronica* 12 / 1: 55-74.

1998a A probabilistic evaluation of Indo-Uralic. In J. C. Salmons and B. D. Joseph (eds), *Nostratic: Sifting the Evidence*. Amsterdam Studies in the Theory and History of Linguistic Science. Series IV. Amsterdam: Benjamins. 153-198.

1998b Some consequences of a new proposal for sub-grouping the IE family. In B. K. Bergen et al. (eds), *Proceedings of the of the Twenty-Fourth annual Meeting of the Berkeley Linguistic Society, February 14-16, 1998*. California: BLS. 32-46

1999a How hard is it to match CVC-roots? *Transactions of the Philological Society* 97: 213-244.

1999b Language classification: scientific and unscientific methods. In B. Sykes (ed.), *The Human Inheritance: Genes, language and Evolution*. Oxford University Press. 45-73.

Ringe, D., Warnow, T. & Taylor, A.

2002 Indo-European and computational cladistics. *Transactions of the Philological Society* 100/1: 59-129.

Rix, H.

1992 *Historische Grammatik des Griechischen. Laut-u. Formenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft [revised edition].

Ross, M. & Durie, M. (eds)

1996 *The Comparative Method Reviewed. Regularity and Irregularity in Langue Change*. Oxford University Press.

Schmidt, J.

1872 *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: H. Böhlau.

Schmitt, R.
1974 Proto-Indo-European culture and archaeology. *Journal of Indo-European Studies* 2/3: 279-88.
1987 Aryans. In E. Yarshater (ed.), *Encyclopaedia Iranica* II /7. London: Rutledge. 684-705.
1989 *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert.
Schmitt, R. (ed.)
1991 *The Bisitun Inscriptions of Darius the Great, Old Persian Text*. London: School of Oriental and African Studies.
Sedlar, J. W.
1980 *India and the Greek World. A Study in the Transmission of Culture*. Totowa (New Jersey): Rowman and Littlefield.
Sihler, A. L.
1995 *New Comparative Grammar of Greek and Latin*. Oxford University Press.
Sims-Williams, N.
1998 The Iranian languages. In A. Giacalone & P. Ramat (eds). 125-53.
Southworth, F. C.
1979 Lexical evidence for early contacts between Indo-Aryan and Dravidian. In M. Deshpande & P. E. Hook (eds), *Aryan and non-Aryan in India*. Ann Arbor MI: Centre for South and Southeast Studies. 191-233.
Steever, S. S.
1993 *Analysis to Synthesis. The development of Complex Verb Morphology in the Dravidian Languages*. Oxford University Press.
1996 Introduction to the Dravidian Languages. In Steever S. S. (ed.), *The Dravidian Languages*. London: Routledge. 1-47.
Sturtevant, E. H.
1933 *Comparative Grammar of the Hittite Language*. Philadelphia: Linguistic Society of America.
1942 *The Indo-Hittite Laryngeals*. Baltimore: Linguistic society of America.
Szemerényi, O. J. L.
1973 La théorie des laryngales de Saussure à Kuryłowicz et à Benveniste. *Bulletin de la Société Linguistique de Paris* 68: 1-25.
1996 *Introduction to Indo-European Linguistics*. Oxford University Press.
Tripathi, D. N. (ed.)
2002/05 *A Discourse on Indo-European Languages and Culture*. Proceedings of the Seminar on The Homeland of Indo European Languages and Culture (Jointly Sponsored by the Indian Council of Historical Research and the Centre for the Study of Indian Tradition). Madhubani, 7-9 January 2002. New Delhi: Manak Publications.

Tucci, G.
1982 *Storia della Filosofia Indiana*. Roma / Bari: Biblioteca Universale Laterza.
Vennemann, T. (ed.)
1989 *The New Sound of Indo-European. Essays on Phonological Reconstruction*. Berlin / New York: Mouton de Gruyter.
Wilbur, T.H.
1977 *The Lautgesetz-Controversy. A Documentation*. Amsterdam: Benjamins.
Winter, W.
1990 Linguistic reconstruction: The scope of historical and comparative linguistics. In E. C. Polomé (ed.). 11-24.
Witzel, M.
1999 Aryan and non-Aryan names in Vedic India: data for the linguistic situation, c. 1900-1500 B.C. In J. Bronkhorst & M. Deshpande (eds), *Aryan and Non Aryan in South Asia*. Harvard Oriental Series III. 337-404.
2000 The home of the Aryans. In A. Hintze & E. Tichy (eds), *Anusantatyai. Festschrift für Johanna Narten zum 70. Geburtstag*. Beihefte Münchener Studien zur Sprachwissenschaft. Beiheft 19: 283-338. Cambridge: Dpt. of Sanskrit and Indian Studies.
Wyatt, W. F. Jr.
1972 *The Greek Prothetic Vowel*. Philological Monographs of the American Philological Association XVII /124.
Zeilfelder, S.
1994/2001 *Archaismus und Ausgliederung. Studien zur sprachlichen Stellung des Hethitischen*. Heidelberg: Winter.

narečii. (Sbornik Otdelenija russkogo jazyka i slovesnosti 1, 2.) Leningrad: Akademija nauk.

Vennemann, Theo
2003 Languages in prehistoric Europe north of the Alps. In: Bammesberger, Alfred & Theo Vennemann in collaboration with Markus Bieswanger & Joachim Hrzega (eds.) *Languages of Prehistoric Europe*, 319–332. Heidelberg. Carl Winter.

Winter, Werner
1978 The distribution of short and long vowels in stems of the type Lithuanian *ė̃sti* : *vèsti* : *mèsti* and OCS *jasti* : *vesti* : *mesti* in Baltic and Slavic languages. In: Fisiak, Jacek (ed.) *Recent Developments in Historical Phonology*, 434–445. (Trends in Linguistics. Studies and Monographs, 4.) The Hague/Paris/New York: Mouton de Gruyter.

Zaliznjak, Anatolij A.
1985 *Ot praslavjanskoj akcentuacii k russkoj*. Moscow: Nauka

Ideology, the Indian Homeland Hypothesis and the Comparative Method

E. Annamalai and S. B. Steever
Yale University

Recent claims about the historical relation of South Asian languages to an Indian homeland have called the Comparative Method of historical linguistics into question. Linguistic comparison of select structural features of Sanskrit and Old Tamil is proposed to determine whether these languages incorporate reference to a 'geographic index' that definitively points to a historical origin in South Asia. The comparison suggests that no such index exists to historically link any of the languages to a particular homeland. Further examination of the issue suggests that the hypothesis of an Indian homeland may owe more to ideological currents of thought than to scientific linguistic analysis.

1. Introduction

1. 1. As is now widely recognized, paradigm shifts (in the sense Kuhn made famous) within fields of academic study may present scholars with an occasion to re-evaluate their field's methods and conclusions to determine whether their traditional content should continue and be incorporated into a new, emerging paradigm, or be set aside as part of the field's history and an object of antiquarian study. Several developments in the social sciences over the past thirty years have offered scholars with a number of such occasions. In particular, the spread of deconstructionism, post-colonialism and post-modernism in the academy has provided scholars with new ways of looking at old phenomena, often equipping them with the tools to challenge received wisdom. While linguistics has not, on the whole, been involved in these particular tides of re-evaluation and self-criticism for reasons

discussed below, recent attempts have nevertheless taken place to apply these new paradigms, particularly post-colonialism, to the subfield of historical linguistics. In doing so, these attempts have called into question a number of previously noncontroversial conclusions about the historical linguistic analysis of the Indo-European languages, specifically its Indo-Aryan branch. Inspired by these new paradigms as well as certain trends in Indian political discourse, several writers have urged that, contrary to traditional consensus, the Indo-Aryan languages originated in a homeland in South Asia rather than arriving there by migration from some geographically external source (with regard to the validity, or otherwise, of this thesis see the chapter by Bryant and Kazanas in this volume and related references). Pursuing this line of reasoning further, others argued that if Indo-Aryan languages are to be related to the Indo-European family at all then India must be the homeland of all Indo-European languages. In rejecting the proposition that Indo-Aryan languages originated outside the subcontinent, some proponents of this sons-of-the-soil hypothesis (Rajaram 1995) went as far as to reject the historical linguistic methodology, specifically the Comparative Method, which had led scholars of past generations to assert a genetic relation between Indo-Aryan and other Indo-European languages.

This new narrative of the development of Indo-Aryan draws much of its inspiration from political ideologies, co-opting academic post-colonial and post-modern ideas to further their claims in the arena of power politics. Sen (2005:67) explicitly links this process to Hindutva as a political ideology during the closing years of the 20th century. This ideology has prompted the re-evaluation of a broad range of interpretations concerning the sources, historical development and nature of the modern Indian state and society. (For a similar instance of co-opting academic history for political purposes, see Jansen's (2000: 482-494) discussion of Meiji Japan.) While this movement appears to offer a post-colonial counterbalance to the claim that imperialist history is written by winners by offering alternative narratives, insofar as it

deals with the issue of language, it has largely skirted or ignored the methods and results of historical linguistics which gave rise to the Indo-European hypothesis two centuries ago and which have tended to confirm it ever since. What results from this re-evaluation is a narrative in which the Comparative Method is at best a caricature and at worst a tool of imperialistic oppression.

1. 2. Viewed from a linguistic perspective, application of the Comparative Method (Ringe 2004) to a set of languages yields at least two kinds of results: it helps us decide when two or more languages are related and, just as importantly, when they are not. Those relationships are expressed in terms of systematic correspondences between the linguistic structures of these languages. One difficulty with the Hindutva-inspired reading of the subcontinent's linguistic prehistory and history is that its 'sons-of-the-soil' argument has focused attention solely on Sanskrit, virtually neglecting comparison with the many other languages and language families whose communities have traditionally occupied the subcontinent. These other families include Dravidian, the Munda branch of Austro-Asiatic and Tibeto-Burman. About these other language families, the sons-of-the-soil movement has been remarkably silent, "miniaturizing" Indian linguistic history, as Sen (2005:69) might characterize it. And when the other languages have been recognized at all, it has sometimes been in terms generally incompatible with linguistic analysis. As Nanda (2003: 81) notes, part of the Hindutva movement has involved the attempt to reclaim "[the] heritage of 'the Aryans' before they were corrupted by miscegenation with the native Dravidians." In popular discourse during the past two hundred years, miscegenation with non-Aryan populations and contact with non-Indo-Aryan languages has lurked in the historical undercurrents as a means by which a pure Sanskrit language and noble Indo-Aryan society have been corrupted, resulting in a state of degeneration that is represented by the diversity of modern Indo-Aryan languages.

Underlying this sons-of-the-soil argument is a geographic fallacy, viz., that the geographic location of a language — what we may call its 'geographic index' — is at all relevant to its potential historical relatedness to other language. Linguistic study of the world's languages over the past two centuries has tended to demonstrate that the geographic locale occupied by a language is neither sufficient nor necessary for classifying that language as to its relatedness, genetic or typological, with a different language or languages. The more fruitful approach, pursued over the past 200 years, has been to ask questions that bear on the properly linguistic, i.e., the structural, aspects of the languages that are being compared, e.g., their phonologies, morphologies and lexicons. Using such correspondences, the application of the Comparative Method to the Indo-Aryan languages continues to uphold the traditional Indo-European hypothesis. Moreover, the new narrative finds itself at a loss to economically explain the systematic correspondences that hold between the Indo-Aryan and Indo-Iranian languages and thence to the remainder of the Indo-European family. Since it is grounded in political and not in linguistic premises, the new narrative is unlikely to shed any light on issues of language and linguistic relatedness; for that, one must ask questions about languages as well as pursue their answers with methodologies appropriate for languages.

We feel that from the perspective of the Indo-Aryan languages, Cardona & Jain (2003b: 32-37) have adequately met the challenge of the Hindutva-inspired narrative to the Comparative Method, and reaffirmed the soundness of both the Method and the Indo-European hypothesis. Both continue to support an interpretation in which the Indo-Aryan language(s) most likely entered South Asia from outside rather than originating autochthonously in the subcontinent. Our contribution to this issue is a modest exercise that tackles the geographic interpretation of language relatedness from another perspective, a perspective that the language situation of the subcontinent offers up. Hindutva ideology appears to hold with a form of linguistic-geographic

determinism, one which claims in effect that the geographic index of a language is a crucial property of that language. This claim supports two corollaries which may be explored through linguistic inquiry. First, if the geographic index of a language is central to the characterization of a language, then it is something that should be directly reflected in the structure of that language independent of any social, cultural or political considerations. Second, if the geographic index of a language is real and relevant to a language's structure, then it should be reflected in much the same way in the structure of all the languages that share the same geographic index. (For if the languages to be compared presented radically distinct linguistic structures, this would call into question the plausibility of a geographic index in the first place.) In the South Asian context, the geographic index thesis effectively claims that the Dravidian, Indo-Aryan, Munda (if we leave out Tibeto-Burman as peripheral to the geography of South Asia) languages are necessarily related because they share indeed the same geographic index. We test this hypothesis by comparing two prominent languages of the subcontinent, Sanskrit and Old Tamil (as is recorded in a few centuries before and after the Common Era), to determine whether their linguistic structures are similar enough to posit the existence of a common geographic index. Even the brief demonstration carried out below reveals stark dissimilarities at many levels of linguistic structure. What similarities are found may be ascribed either to the influence of linguistic typology, which inspection of the world's languages shows is not geographically determined, or to linguistic convergence in the formation of a linguistic area (Masica 1976), which can be documented through appeal to the historical records as well as historical linguistic techniques. Thus, those features of linguistic structures that are areally shared by both Sanskrit and Old Tamil are not indicative of geographical index for the simple reason that they can be empirically shown to be historical differentials having been acquired through contact. These features do not manifest when there is no bilingualism among the speakers. Geographic

index is independent of bilingualism. This comparison strongly suggests that there is no geographic index which is connected to or exerts an influence over a language, and that appeal to such an index has no utility in historical linguistic discourse. This in turn vitiates the criticism which the Comparative Method has recently received from ideological quarters.

We hope to show below that a geographic index is not a sufficient criterion for genetic relatedness because the Dravidian and Munda languages, even though they have shared the subcontinent with Indo-Aryan languages from prehistoric times, and thus presumably have the same geographic index, are demonstrably unrelated to Indo-Aryan. Following the conventional viewpoint, we distinguish these three language families on a genetic basis, but allow for their typological convergence over time to form a linguistic area in South Asia. Yet, despite the formation and evolution of a linguistic area in South Asia, what application of the Comparative Method to the languages of this region has demonstrated over the past 150 years is that the Dravidian languages belong together, and that they are genetically distinct from Indo-Aryan and Munda, which are also genetically distinct from each other. Thus the geographic index of a language, to the extent it can be said to exist, proves to be irrelevant to linguistic issues such as historical origin, whereas systematic structural relations between two or more languages prove to be decisive in issues of relatedness.

Ours is not of course a novel exercise but one that was first carried out in the first half of the 19th century. In 1801 Colebrooke classified languages such as Tamil and Bengali, Telugu and Hindi as linguistically related because they were "cultivated," i.e., they had established writing systems and literatures. Linguistic science, as it evolved through that period, began to separate language and writing systems, on the one hand, and language and literary composition, on the other, as distinct cultural systems to be studied separately. By 1857, when Bishop Caldwell published his comparative grammar of the Dravidian languages, he not only

established the Dravidian languages as a family distinct from Indo-European on structural and historical grounds, but also distinguished the Munda subgroup of Austro-Asiatic as separate from both. Even during that era, claims of linguistic relatedness were being used to bolster political ideologies long after linguists had abandoned the claim that all cultivated languages of the subcontinent were related (Steever 1999). While we are confident that our demonstration can be replicated with the appropriate comparisons with Munda and Tibeto-Burman languages, our expertise limits us to the Dravidian family. However, the rehearsal of this demonstration, first undertaken 150 years ago, provides us with a way to critically view the current post-colonial critique of the Comparative Method.

For the comparison, we draw from readily available descriptions of Sanskrit and Old Tamil. For Sanskrit, we have occasion to refer to Burrow (1977), Masica (1999), Cardona (2003) and Jamison (2004); for Tamil, to Lehmann (1994) and Steever (2004). The comparison will be brief; readers are referred to these publications for greater detail.

2. Phonology

Old Tamil (OT) has a single set of stops which are phonemically distinguished by six places of articulation but not by manner of articulation, e.g., voiced versus voiceless. OT distinguishes labial, dental, alveolar, retroflex, palatal and velar series, with the alveolar and retroflex stops having a more restricted distribution than the others. By contrast, Sanskrit (Skt.) distinguishes five places of articulation, labial, dental, retroflex, palatal and velar. Proto-Indo-European (PIE) labiovelars merged with plain velars in Sanskrit, as happened in other *satem* languages. While this may appear to suggest that the predecessor of Sanskrit had six places of articulation, making it more comparable with Old Tamil, that impression is incorrect because there is no series of retroflex stops in early Indo-Iranian, much less Indo-European, to which the

Sanskrit retroflex series may be traced. Sanskrit acquires a retroflex place of articulation through language-internal developments (*s > *ẓ -/ {r, u, k, i}) which are gradually later extended to other environments through regressive assimilation in consonant clusters; etymologically, the retroflex stops generally descend from dentals. Famously, Proto-Dravidian (PDr) and by extension Old Tamil have a retroflex place of articulation in their consonant systems, but in Old Tamil and the rest of Dravidian the dental and retroflex series are etymologically distinct and do not alternate. In Sanskrit velars and palatals alternate (reflecting an earlier palatalization of velars following front vowels (and *y*) whose conditioning was subsequently lost), as in reduplicative forms such as *ca-kaara* (√ *kṛ*) 'do'. OT does not evince such alternations in the synchronic grammar of the language. The plain voiceless, voiced and voiced aspirated (breathy voice) series of stops of Sanskrit are inherited from PIE while the voiceless aspirated series is innovated. As noted above, Old Tamil does not distinguish stops according to manner of articulation at all.

The cardinal vowels of Sanskrit (Jamison 2004: 676), *a*, *i*, *u*, distinguish length while the mid vowels *e* and *o* are always long. Old Tamil distinguishes short and long versions in all five primary vowel qualities, *a*, *i*, *u*, *e*, *o*. Sanskrit has two diphthongs, *ai* and *au*, both of which are long. It also has vocalic liquids *ṛ* and *ḷ*, which lack counterparts in Old Tamil. Sanskrit possesses an extensive system of vowel alternations, ranging from zero-grade, through full grade (guNa) to extended grade (vRddhi) that is abundantly reflected in morphological alternations in the language. Old Tamil, by contrast, contains only a handful of lexically determined vowel length alternations, e.g. *taa-* vs. *ta*(*ra*) 'give to you or me'.

In the area of supra-segmentals, PIE and Vedic Sanskrit had a pitch accent system (Jamison 2004: 680) which had grammatical and potentially lexical functions; neither Old Tamil nor the rest of Dravidian have anything comparable. Thus the phonology of the two languages is distinct, whereas comparisons with such languages as Ancient Greek and Avestan reveal systematic correspondences with Sanskrit but not Old Tamil.

3. Morphology

In general, Sanskrit morphology exhibits a synthetic character while Old Tamil is predominantly agglutinative although the latter does show signs of fusion (see Lehmann 1994). In fact, it may be the agglutinative character of Dravidian morphology that restrains the development of the complex morpho-phonological alternations that are so characteristic of internal sandhi in Sanskrit. Both Sanskrit and Old Tamil are predominantly suffixal in nature. Sanskrit appears to permit just three prefixes; some forms in Old Tamil that appear to be prefixes, such as the first element in the demonstrative series, may well be the first part of a compound nouns. Sanskrit permits tmesis, OT does not.

<u>Parts of speech</u>. Old Tamil distinguishes primarily between free forms (words) and bound forms (postclitic particles). The free forms include nouns and verbs, which are distinguished according to the grammatical categories they encode and the inflections they bear. As to the question of additional parts of speech, scholarly consensus is divided. Some (Lehmann 1994) admit a further though highly restricted category of adjectives while others see them as defective forms of nouns or verbs. Sanskrit (Jamison 2004:682) distinguishes between inflected and a small class of uninflected words. The latter class includes such important operators as negative markers and conjunctions. The former class includes nominal and verbal phrases. Adjectives in Sanskrit are far more formally developed than their notional counterparts in Old Tamil; whether basic adjectives or participles derived from verbs, they participate in the nominal inflectional categories of Sanskrit. Further, Sanskrit adjectives morphologically manifest comparative and superlative forms. By contrast the so-called adjectives of Old Tamil do not participate in the nominal inflections of the language. There are no comparative or superlative forms; comparison is

encoded periphrastically so that a phrase of the type: 'He is taller than his brother' is rendered as: 'In comparison with his brother, he is tall'. Furthermore, adjectives in Old Tamil cannot serve as hosts to postclitic particles.

Nouns. Sanskrit has a number system that distinguishes singular, dual and plural numbers while Old Tamil distinguishes only between singular and plural. Sanskrit recognizes three genders: masculine, feminine and neuter. While feminine gender is signaled primarily by means of derivation, masculine and neuter are signaled by inflection. Gender in Old Tamil is basically a system of natural gender that distinguishes primarily between animate and inanimate (Steever 2004: 1020). It should be noted that predicate nominals in Old Tamil, but apparently not Sanskrit, may inflect for personal endings that agree in person, number and gender with their subjects. Sanskrit has eight cases: nominative, accusative, instrumental, dative, ablative, genitive, locative and vocative. These cases are indicated by means of suffixes. Old Tamil grammarians recognize eight cases, as well: nominative, accusative, dative, instrumental, equative, genitive, locative and vocative. The Old Tamil case markers are often supplemented by the use of postpositions, particularly when one wishes to further specify a spatial relation, of the type of, e.g.: 'from inside of'. Despite the similarity in nomenclature, however, the functions of the cases are dissimilar in the two languages. For example, certain instances of the dative case in Old Tamil appear to have subject-coding properties, giving rise to dative-subject constructions (Steever 2004:1035), which appear to lack a counterpart in Sanskrit.

Pronouns. Old Tamil distinguishes between inclusive ('includes the speaker') and exclusive ('excludes the speaker') first person plural pronouns; Sanskrit does not. As noted above, Sanskrit has dual pronouns distinct from singular and plural whereas Old Tamil lacks a dual altogether. Other Indo-European languages, such as Ancient Greek, preserve a dual number.

Verbs. The set of verbal bases in Old Tamil, as in most of Dravidian, is a closed one; consequently, this inhibits the possibility of extensive verbal derivation. Old Tamil verbs distinguish primarily two tenses: past and non-past while Sanskrit distinguishes more (Jamison 2004:688): five in Rig Vedic Sanskrit and four in post-Rig Vedic Sanskrit. Negation is a matter of verb inflection in Old Tamil, i.e., the language possesses a separate negative conjugation, while negation in Sanskrit is expressed syntactically through the juxtaposition of independent particles such as *na* to verb forms. Both Sanskrit and Old Tamil distinguish between finite and nonfinite verb forms. However, their distribution appears to be governed by different principles. The distribution of finite verbs in Old Tamil is governed by a rule made explicit in Steever (1988), while to the best of our knowledge a similar rule, if applicable, has yet to be worked out for Sanskrit or Indo-Aryan. It should be noted that predicate nominals in Old Tamil have the same distribution as finite verbs; in fact, many predicate nominals bear personal endings that show person-number-gender agreement with the subject, a phenomenon well attested in other Dravidian languages, but absent from Sanskrit. Old Tamil includes a set of constructions called serial verb constructions (Steever 1988), in which the members of a verb-verb compound both contain personal endings that are congruent. Sanskrit lacks such constructions even though certain daughter languages such as Urdu, Magahai and Gujarati have developed them (Cardona & Jain 2003: 290, 323, 511, 68).

Clitics. Clitics in both Old Tamil and Sanskrit appear after their hosts; as a result, neither can begin a sentence. In Sanskrit, following Wackernagel's Law, clitics typically attach after the initial accented word of the sentence (Jamison 2004:696) whereas Dravidian clitics are not restricted to such a position, but may

occur virtually anywhere in a sentence its host appears. There are restrictions on individual clitics; for example, the OT clitic *-um* 'and, all' may not combine with a finite verb whereas Skt. *ca* 'and' appears to exhibit no such restriction.

Syntax. Both Old Tamil and Sanskrit appear to favor SOV (Subject-Object-Verb) word-order. In Sanskrit this is the default word order in prose (Jamison 2004:695), although the case system obviates the need for a rigid word order to signal grammatical relations. Because both are inflectional languages, Sanskrit and Old Tamil are able to exploit the potential of word-order variations in poetry. This is conspicuous in Old Tamil, the bulk of whose corpus is poetic in nature. What prose there is tends to appear in the inscriptional corpus (see Mahadevan 2003); there SOV word order predominates. Beyond superficial word order, Old Tamil rigorously exhibits a variety of correlations associated with SOV word-order: it has postpostions, not prepositions; genitives precede rather than follow their heads; main verbs precede rather than follow auxiliary verbs; syntactic structures branch to the left, not the right; etc.

Agreement. Both Old Tamil and Sanskrit are nominative-accusative languages. Additionally, Old Tamil has dative-subject constructions in which a noun phrase in the dative manifests some subject-coding properties, such as the ability to antecede pronouns, but lacks others, such as the ability to trigger subject-verb agreement. Sanskrit appears to lack such non-nominative subjects, although later Indo-Aryan certainly develops constructions that resemble them, as in Bangla, or the split ergative-nominative system in Hindi.

Lexicon. Comparison of the Dravidian Etymological Dictionary with Turner's dictionary of Indo-Aryan reveals two distinct lexicons. Comparison of kinship terms, number systems and other basic elements of vocabulary give rise to no systematic

correspondences between Old Tamil and Sanskrit. Close examination of the vocabulary in the Old Tamil classics reveals a society which is well-integrated with its environment: the theory of 'tiNai', or landscapes, shows that early Tamils were well aware of the flora and fauna around their South Indian homeland. This sense of possession of place is not as evident in Indo-Aryan, whose vocabulary betrays traces of a more northerly source, particularly in the words for various trees. Correspondences between the vocabulary of Sanskrit and that of other Indo-European languages are far stronger than those between Sanskrit and Old Tamil. Comparing Old Tamil with its medieval and modern successors reveals, naturally enough, that Tamil has borrowed lexical items from Indo-Aryan (and vice versa), and, up until the early 20th century, the proportion of such borrowings steadily increased. But even though the earliest stages of the language show borrowings from Indo-Aryan, they are meager and unsystematic. Comparisons of Sanskrit vocabulary with the vocabularies of other Indo-European languages confirms numerous lexical correspondences, which cannot be secured when comparing Sanskrit and Old Tamil

4. Conclusion

This brief comparison shows that Sanskrit and Old Tamil are distinct languages at every level of their structure. If we cast our net wider to encompass more languages, further comparison will show that Sanskrit shares more in common with the structures of Avestan, Ancient Greek, Latin and Hittite than it does with Tamil or Sora. By the same token, Old Tamil will be seen to share more in common with Telugu, Kannada, Gondi and Kurux than it does with Dhivehi or Bangla. It is these patterns of structural correspondence that give rise to the Indo-European and Dravidian families, respectively. Thus possession of the same geographic index is no guarantee of linguistic relatedness. And if that is the case, then one pillar of the criticism against the Comparative

Method is removed. Accordingly, the claim that Indo-Aryan languages necessarily originated in a South Asian homeland receives no confirmation from a linguistic analysis and comparison of their linguistic structures with those of other languages, notably Dravidian in South Asian and Indo-European languages in West Asia and Europe. The comparative method is thus reaffirmed in two ways. It permits us to reconstruct a fragment of Proto-Dravidian; furthermore, once we strip away the accretions of borrowing, purely internal developments and the like, we can easily identify Indo-Aryan as a language subgroup that belongs to a larger Indo-Iranian group which, in turn, belongs to the Indo-European family of languages.

As observed in the introduction, the discipline of linguistics, including historical linguistics, has not participated in the critical re-evaluation occasioned by the spread of such paradigms as deconstructionism, post-colonialism or post-modernism. Accordingly, the brute force application of the Hindutva-inspired post-colonial critique of the Comparative Method, rebutted above, will fail to persuade many linguists. Does this mean that linguistics is immune from criticism? Far from it. Any comprehensive history of linguistics, particularly historical linguistics, over the past 200 years will reveal that the Comparative Method has survived numerous paradigm shifts and challenges to its validity during that time. In its earliest phase, the Comparative Method emerged in the late 18^{th} and early 19^{th} centuries and gathered confirmation through an impressive body of evidence. It was challenged by the Neo-grammarian hypothesis of regular sound change in the second half of the 19^{th} century and emerged from that test stronger than before. Going forward, the Comparative Method also passed through Saussure's revolution of distinguishing between synchronic and diachronic treatments of language, contributing to that revolution in no small part. It then passed through the paradigms of American descriptivism, European structuralism and generative grammar, gaining strength from each challenge. The survival and growth of the Comparative

Method through these various paradigm shifts stands as testimony to its continued utility in historical linguistic analysis. Given the robustness the Comparative Method has displayed through all these major paradigm shifts in the history of linguistics, it seems unlikely that post-colonialism will be able to strike a decisive blow against this time-tested methodology. (After all, post-colonial historians have not given up the use of documents in writing history.) The reason for this is that post-colonialism, as with post-modernism and deconstructionism, appears to concern itself primarily with ideology and not with a scientific analysis of language structure and change (for an alternative point of view on this topic please consult the chapter by Bryant in this volume). Ideology, in the Marxist sense, is a theory that is not based on the will to transform reality and therefore becomes an abstract construct to justify a corrupt power structure. Linguistic analysis, on the other hand, aims to transform our understanding of the nature of language by revealing relations, structures and principles that were not seen hitherto. A deconstructionist might reply that everything is about power, including language, and that even the rarefied study of linguistics is not exempt from its siren call. But if everything is about power, then nothing is about power because we can only recognize power by contrasting situations in which its exercise is vital or necessary with those in which it is not. Given the speed with which linguistic theory and practice have developed over the past 50 years, the Comparative Method will almost certainly be tested again by new, inchoate linguistic paradigms, but, we believe, ideology will have little impact on its development.

References

Burrow, Thomas
1977 *The Sanskrit language*. London: Faber and Faber.
Cardona, George
2003 Sanskrit. In Cardona & Jain 2003a (eds). 104-160.
Cardona, George & Dhanesh Jain (eds).
2003a *The Indo-Aryan Languages*. Routledge.
2003b Introduction. In Cardona and Jain 2003a (eds). 1-45.
Jamison, Stephanie
2004 Sanskrit. In Roger Woodard (ed.), *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge: Cambridge University Press. 673-699.
Jansen, Marcus B.
2000 *The making of modern Japan*. Cambridge: Harvard University Press.
Lehmannn Thomas
1994 *Grammatik des Alttamil*. Stuttgart: Steiner.
Mahadevan, Iravatham
2003 *Early Tamil epigraphy: From earliest times to the sixth century A.D.* Chennai: Cre-A.
Masica, Colin
1976 *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.
1991 *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
Nanda, Meera
2003 *Postmodernism and religious fundamentalism: A scientific rebuttal to Hindu science*. Pondicherry: Navayana.
Rajaram, Navaratna S.
1995 *The Politics of History: Aryan Invasion Theory and Subversion of History*. New Delhi: Voice of India.
Ringe, Don
2004 Reconstructed ancient languages. In Roger Woodard (ed.), *The Cambridge Encyclopedia of the World's Ancient Languages*, pp. 1112-1128.
Sen, Amartya
2005 *The Argumentative Indian: Writings on Indian History, Culture and Identity*. New York: Farrar, Straus and Giroux.
Steever, Sanford
1988 *The serial verb formation in the Dravidian languages*. Delhi: Motilal Banarsidas.

1999 Review of T. R. Trautmann.1997: *Aryans and British India*. Berkeley: University of California Press. xiv + 260. *Language* 75/1: 147-48.
2005 Old Tamil. In Roger Woodard (ed), *The Cambridge Encyclopedia of the World's Ancient Languages*. Cambridge: Cambridge University Press. 1015-1040.
Trautmann, Thomas
1997 *Aryans and British India*. Berkeley: University of California Press. xiv + 260.

The Indo-Aryan Migration Debate

Edwin Bryant
Rutgers, New Jersey

The Indo-Aryan Invasion debate has reached a point of stagnation and, in the absence of compelling new data or interpretative models, both sides remain entrenched in their respective convictions as to whether the Indo-Aryan of South Asia were originally indigents or migrants into the subcontinent. Here I shall examine some of the ideological trappings motivating the production of knowledge around this issue, as well as some of the more prominent aspects of the relevant data. In accordance with one of the sub-themes of this volume, this chapter will probe for circularity underpinning some of the evidence. Additionally, given the challenges to the normative model of Indo-European reconstruction from some other chapters of this volume, this chapter will underscore the fact that the South Asian linguistic evidence and all the controversy it has engendered is derivative from this model. Any reconsideration of normative Indo-European paradigms will thus hold significant repercussions on the proto-history of India.

1. The Indo-European homeland problem

1. 1. The IE homeland problem has tantalized scholars for almost two centuries, and resulted in a massive production of knowledge dedicating itself to reconstructing a PIE language, locate an original physical homeland where it was spoken, conjecture on the social and cultural life of the proto-speakers, and chart their historical emigrations from this geographical point of origin. While certain conventionalist perspectives on PIE may find such 'realist' reifications misguided and passé, the notion of IE origins has remained the foundation stone of proto-historic reconstruction for a huge swath of human civilizations. Conceptualized and pursued, often brilliantly, by Western scholars motivated initially

by an interest in the origins of western civilization, this academic undertaking has necessarily caused scholars to attempt to reconstruct the proto-histories of non-Western countries that happen to partake of the Indo-European language family. The topic of this paper focuses on the repercussions of Indo-European studies on scholarship pertaining to one such group: the Indo-Aryans of ancient India.

As a result of the family tree model of IE reconstruction, since around the middle of the 19th century, text books on early South Asian history and civilization have outlined some version of events whereby an IE sub-branch called the Indo-Aryans entered into the Indian subcontinent sometime in the second millennium B.C.E. from some external point of origin, and introduced the Vedic culture into South Asia. This version of events has most commonly been referred to as the 'Indo-Aryan Invasion theory', reconfigured more recently by its subscribers, into the 'Indo-Aryan Migration theory'. Although not without its critics since its conception, this theory has been vociferously challenged over the last couple of decades, by primarily India, but also some Western scholars, on the grounds that there is no convincing evidence to posit such an invasion/migration, and the issue was one of the most hotly contested debates in Indology in the 90's. The debate has died down in Western academic circles, somewhat, recently, not really because it has been resolved decisively, but, in part, because scholars became exhausted with the polemical and emotional tenor of the discussion, and the missionary zeal with which the opposing views were pursued. Moreover, few scholars command the requisite knowledge (or, perhaps, more accurately, have the time) to form an independent opinion based on personal analysis of the enormous amount of multi-faceted, multi-disciplinary data involved.

The issue, however, remains very much alive in India at the highest academic levels[1] and amongst the Indian diaspora[2]. As is unavoidably evident to anyone in South Asian area studies, the reconstruction of the history of the Vedic and pre-Vedic speakers is not just of relevance to the greater field of Indo-European

studies, but also of intense interest to many scholars in the present post-colonial context of ancient Indian historical reconstruction. Indeed, many critics of the Migrationist view from within the academy have been Indian archaeologists (we need not refer to the plethora of publications surfacing on the matter outside of academia in India and the diaspora). Opposition to the theory runs the gamut from an agnostic stance which simply finds all the Invasionist / Migrationist circular, faulty, or inconclusive, to an Indigenist stance which claims that the Indo-Aryans were indigenous to the subcontinent (a position generally disinterested in the greater IE language connections), to a more assertive Out-of-India position, which retains the IE family tree (or derivative wave) model and sets out to reconstruct some of the dates to support an emigration of non-Sanskritic IE languages to the West from an Indian *Urheimat* (for which see Kazanas in this volume). For the purpose of simplicity, I will categorize all these positions in this paper as the 'Indigenist' view irrespective of the variations noted above. Since few speak of 'invasions' anymore, I will refer to the more normative view as the 'Migrationist' view.

Knowledge, of course, is never produced in a social vacuum, and the problem of IE origins has vexed scholars in both India and the West for well over a century, touching all kinds of nerves in both academic and political communities. This enormous intellectual enterprise, which has a well-known history of ideological appropriation in Europe, has, especially over the last two decades, found its way into political discourse in India, where it continues to provoke all kinds of emotionalized, polemical and *ad hominem* controversy. While it remains a fascinating problem, the discussion has unfortunately become increasingly and seemingly inescapably politicized, emotional, and strident, and labels of 'Hindu nationalist,' 'western neo-colonialist,' 'Marxist secularist' and other such simplistic and derogatory stereotypes are frequently bandied about. One does not have to lurk for long on any one of a number of South Asian list-serves, or consider the California text book controversy[3], to soon encounter the issue and the diametrically opposed views that it generates, views that are

all-too-often articulated in condescending, hostile, and emotional tones. Clearly, there is a lot more at stake, here, than the existence or absence of innocuous horse bones in the Indus Valley, or the different possible ways of accounting for the curious non-IE linguistic features in the *Rgveda*.

1. 2. Having published a monograph as well as an edited volume discussing the range of data involved in the debate over Indo-Aryan origins, as well as some of the socio-politico-religious contexts of its interpretations over the last two centuries (Bryant 2001 & 2005), my role in this volume is to sieve through my own work and identify some of the more conspicuous elements in the history of the debate. I wish to state, first, that I have nothing new to add to the analysis I have already directed to the issue, which I feel has reached a state of stagnation in the absence of compelling new evidence. I have agreed to tailor some of the highlights of my already published work, which was primarily directed to Indologists, for this volume, since is aimed more specifically at Indo-European scholars, who may be less acutely aware of the intense controversy this issue has aroused in India. However, since Marcantonio (in her *Introduction* as well as in her contribution to the volume) has drawn attention to circularity that she holds envelops some of the primary aspects of historical IE reconstruction, I will revisit my previous research from the perspective of the circularity that underpins much of the dates pertaining to Indo-Aryan migrations into India. One thing that can be stated forthwith, is that the entire issue of Indo-Aryan migrations is a *derivative* consequence of the 'family tree' (or its modification as a 'wave') presuppositions of historical linguistics (i.e. that there was some sort of a PIE language community or cluster of communities inhabiting some sort of circumscribed area, and that derivative or daughter languages did splinter off from this area to other areas such as the Indian subcontinent where non-IE languages had previously been spoken). We might well note at the outset that were there to be any degree of consensus in mainstream IE circles that normative and foundational IE reconstructive

methods were based on inadequate or faulty presuppositions and methods, there is nothing in Indian archaeology that demands or actually in any way supports a migration of peoples into the Indian subcontinent during even the broadest of time frames associated with the post-PIE dispersions. Since the interpretation of data relevant to historical reconstruction always takes place within a cultural context, and given the exhausting emotionality that the issue engenders in Indological circles, in addition to attempting to draw attention to the malleable and often circular nature of much of the evidence itself, I will suggest that there are two primary ideological reasons why many (but, let us be clear, by no means all) Indian scholars have seen fit to revisit or question the theory of the external origins of the Indo-Aryans. The first is suspicion of the motives of 19^{th} century scholars who pieced together the theory in the first place, and the second is the imperatives of Hindu nationalism. I can do no more, here, than touch upon some of the highlights of the entire controversy, but will attempt to do so with a view to how and why the whole enterprise has provoked reaction from the other side of the fence, so to speak, from the perspectives of certain Indian intellectuals. The 'discovery' of the Sanskrit language and literatures by Europeans in the late eighteenth century had enormous corollaries for the absolute authority of Abrahamic historical genealogy pertaining to the genesis of the human race. If accepted at face value, the Sanskrit material, as scholars such as Sir William Jones well knew, threatened to subvert the absolute authority of Mosaic history, a prospect he and many scholars found unacceptable. Accordingly, there were efforts made, either to make the Hindu chronology conform to that of *Genesis*, which Jones and others struggled to do, or, to herald the newly discovered material as evidence to discredit Biblical narrative as other segments of the intelligentsia such as Voltaire set out to do. Either way, the Sanskrit material was initially filtered through the position European scholars took on the bible, in the late eighteenth and early nineteenth century. But even well after the tower of Babel was no longer in the picture, biblical sensitivities were replaced by colonial ones, and there was a very

cool reception in some circles to the "late Prof Max Müller [who had] blurted forth to a not over-grateful world the news that we and our revolted sepoys were of the same human family" (Legge 1902: 710). One need only quote Müller (1883: 28) himself to get a sense of the mood of the day:

> They would not have it, they would not believe that there could be any community of origin between the people of Athens and Rome, and the so-called Niggers of India. The classical scholars scouted the idea, and I still remember the time, when I was a student at Leipzig and began to study Sanskrit, with what contempt any remarks on Sanskrit or comparative grammar were treated by my teachers...No one ever was for a time so completely laughed down as Professor Bopp, when he first published his Comparative Grammar of Sanskrit, Zend, Greek, Latin and Gothic. All hands were against him....How violent a shock was given by the discovery of Sanskrit to prejudices most deeply engrained in the mind of every educated man. The most absurd arguments found favor for a time, if they could only furnish a loophole by which to escape the unpleasant conclusion that Greek and Latin were of the same kith and kin as the language of the black inhabitants of India

Nonetheless, the discoveries of the fledgling science of historical linguistics reality proved to be unavoidable, and colonial sensitivities had to find other ways of negotiating with the corollaries of the IE language connection. Before long, some colonial administrators found ingenious ways of reconfiguring Aryanism in politically advantageous ways. As Maine (1875: 18-19) puts it: "The government of India by the English has been rendered appreciably easier by the discoveries which have brought home to the educated of both races the common Aryan parentage of Englishman and Hindoo ... in thus honoring our Aryan forerunners in India we shall both honor ourselves and make the most direct and effective appeal to Indian loyalty". Parallel statements were made by certain Christian evangelists who also found similar advantages in celebrating Aryan kinship (e.g. Hastie's (1882) appeal to the Indians to shake off their servitude to the IE gods and embrace Christ, as the European side of the family

had done many centuries earlier). When one factor in the direction Aryanism took in Germany, which was to some extent fortified with dubious readings of certain Aryan-related passages in the Vedic texts, one can perhaps be more sympathetic as to why a good number of Indian scholars view the whole production of Aryanism as ideologically tainted from its inception.

We need add to this the fact that, even after two centuries of intense speculation, scholars still hardly agree on even the most basic details of the Indo-Europeans – their culture, their origins, or almost anything at all. One need only consider a few of the current theories in the field: the archaeological evidence, for example, has been configured on the one hand by the Gimbutas school to promote a mounted, nomadic IE warrior from the 4th millennium BCE Steppes, and on the other by Lord Renfrew and his school to argue for a peaceful sedentary IE agriculturalist from 7th millennium BCE Anatolia. Meanwhile, linguistic data suggests on the one hand a warm exotic Near Eastern homeland with elephants and monkeys to the linguists Gamkrelidze & Ivanov (1983a & b), but a homeland much further East, in Bactria to their translator, Nichols (1997, 1998).

In short, then, given the religious, colonial and nationalistic contexts of almost the entire history of IE studies, and given the remarkable lack of consensus regarding the point of origin or even the basic cultural characteristics of the Indo-Europeans despite over two hundred years of intense research, it is not terribly surprising that, in the sensitive post-colonial climate of recent years, a significant number of Indian scholars are going to wonder why on earth they should feel obliged to retain theories pertaining to the origins of the South Asian side of the family that emerged from such contexts during the 19th century. As one cynic put it:

> Instead of letting us know definitely and precisely where the so-called original home of the Aryans lay, they drag us into a maze of conjectures clouded by the haze of presumptions. The whole subject of the Aryan problem is a farrago of linguistic speculations or archaeological imaginations complicated by racial prejudices and chauvinistic xenophobia. It is high time we extricate ourselves from this chaos of bias and belief (Prakash 1996: 44).

2. Modern Indian historiography

Having noted something about the European ideological matrix of scholarship on this issue, we must now turn our attention to the communal element in modern Indian historiography. The present volatile situation in India has made western, and many Indian, scholars particularly concerned about the repercussions of communal interpretations of Vedic history in that country. Once can hardly avoid noticing the fact that the vigor with which the whole issue was revisited in the 1980s and 90s overlapped with the ascendancy of Hindu nationalism, and while one must be cautious of lumping all Migrationists into one simplistic ideologically motivated category, one cannot also deny a certain amount of unscholarly opportunism that has motivated, or at least taken advantage of, some of this genre of revisionism. Therefore, let us get the ugliest face of nationalistic Hindu revisionism out on the table. Perhaps the best illustration of extremist nationalist discourse can be found in the writings of Gowalkar (1947:53), the ideological guru of the '*Hindutva*'[4] rightwing on the Aryan issue:

> It was the wily foreigner, the Britisher, who, to achieve his ulterior imperialistic motives, set afloat all such mischievous notions among our people so that the sense of patriotism and duty towards the integrated personality of our motherland was corroded. He carried on the insidious propaganda that we were never one nation, that we were never the children of the soil *but mere upstarts having no better claims than the foreign hordes of Muslim* or the British over this country.

The concerns, here, are evident: if the Vedic Indo-Aryans are construed as being indigenous to India, then the 'Vedic Civilization' and all that developed from it can be perceived as 'truly Indian' and all subsequent cultural groups known to have immigrated into India can be depicted as 'Others.' From this

perspective, if Hindus can be construed as *deshi*, primordial sons of the soil, then their claim to defining the ethos of a modern nation state is enhanced. Thus, this brand of Hindu nationalism, which seems determined to alienate the Muslim and Christian communities on the grounds of their lack of indigenous historical pedigree, is obliged to refute the Aryan invasion theory. Naturally, such claims can only be made at the expense of non-Hindu communities perceived as 'Other.' Again, to quote the most extreme articulation this position has received, even in right wing ideological circles (Gowalkar 1947: 55-57)

> The non-Hindu people in Hindusthan must either adopt the Hindu culture and language, must learn to respect and hold in reverence the Hindu religion, must entertain no idea but those of glorification of the Hindu race and culture...or may stay in the country, wholly subordinated to the Hindu nation, claiming nothing, deserving no privileges...not even citizen's rights

Having said this, however, it seems fair to add that although the promotion of Indigenism is undoubtedly extremely important to notions of identity and to the politics of legitimacy amongst certain Hindu nationalists, such concerns are not representative of all the scholars who have supported this point of view. Unfortunately, the whole Indigenist position is often simplistically stereotyped, and demonized, both in India and the West, as a discourse exclusively determined by such agendas. This simply provokes counter-stereotypification in kind, and Migrationists in India are labeled 'secular Marxists' or 'colonial stooges' by their detractors and accused of manipulating the data in support of their agenda, namely, an insistence that India is a hodge-podge of immigrants within which no group has Indigenous pedigree, or a greater claim to cultural hegemony, than any other. Although there are doubtlessly nationalistic and, in some quarters, communal agendas lurking behind some of this scholarship, other concerns also motivate such reconsideration of history, namely, the desire of many Indian scholars to reclaim control over the reconstruction of the religious and cultural history of their country from the legacy

of imperial and colonial scholarship. As I suggested in the first part of this paper, and as few would deny, there are reasons to question the context of 19th century scholarship on the issue (even as harangues against colonial Indology are by now becoming somewhat tiresome). I have, in the past, expressed concern at what I have termed a type of Indological McCarthyism creeping into areas of western, as well as certain Indian, academic circles, whereby *anyone* reconsidering the status quo of Indo-Aryan origins is instantly and *a priori* dubbed a nationalist, communalist or, even worse, a Nazi (Bryant 2001). Not all Indigenists are necessarily interested in the construction of notions of Hindu Aryan greatness or, with some troubling exceptions, in the promotion of communal agendas.

There is a tendency of participants in this debate in Indology circles to be highly selective in the appropriation of the data involved in Indo-Aryan origins. Attempts to approach the ideals of objective scholarship by discussing the valid problems that have been pointed out on both sides, can be hard to come by, even if one feels one version of events better accounts for the entirety of the evidence, There is all-too-often a mocking and condescending tone adopted by disputants of this issue, and a tendency of simply highlighting and ridiculing the most outlandish aspects of an opponents arguments while ignoring any coherent points that might counter one's own position. There is also a tendency to immediately resort to tactics of casting *ad hominem* aspersions – lumping those who hold an opposing view point, whether Migrationist or Indigenist, into simplistic categories and branding them 'nationalist,' 'neo-colonialist,' 'traditionalist,' 'Marxist,' or some other such hasty and sometimes unwarranted category. The hostility, zealousness and *ad hominem* sarcasm seeping from the pens of participants in this debate (from both sides of the fence) when referring (increasingly by name) to those holding opposing views is blatant. Consequently, the academic value of much of the exchange – on all sides – has been singed by the emotional temperature such issues ignite. The result, especially in India, is an almost complete lack of communication between

two mutually antagonistic and angry camps, and intransigent, cavalier, selective and often grossly inaccurate generalizations of opposing views are bandied about on both sides of the issue. Thus, an entire fascinating field of study has become inextricably linked with ideology and the politics of representation to the point where it is almost impossible to have a rational and objective conversation on the origin of the Indo-Aryans in India without becoming associated with the ideologies that are immediately correlated with pro- or contra- stances on the issue.

In any event, while the relevance and indeed, centrality, of Indigenism to the politics of legitimacy and identity amongst certain Hindu nationalists cannot be denied, such ideological agendas are by no means representative of all the scholars who have supported the Indigenist point of view. There are many Indian scholars who simply do not feel the data supports a Migration thesis; they cannot all be tarred and feather with the same simplistic 'communal' brush. In short, there are three sets of motives fueling much Indigenist revisionism: suspicion of the ideological agendas of European scholars who pieced together the theory in the first place, the imperatives of renaissance Hindu nationalism in some quarters; and the extreme malleability and circularity of much of the evidence itself. It is to this malleability of evidence that we now turn.

The suspicion of the European context of this production of knowledge by Indian intellectuals was not assuaged when much of the data brought forward in support of the Indo-Aryan Migration hypothesis was brought under scrutiny. One might note, and it seems fair to generalize here even as there have been exceptions, that Indian scholars were and still are primarily interested in the entire IE issue only in so far as it touched on the origins of their own Vedic forefathers, rather than on the entirety of the IE family. Access to the cutting edge work of historical linguists in Europe still remains hard to come by for scholars in the few departments with where IE linguistics is taught in India (and, therefore, with a few notable exceptions, they have contributed much less to the field of IE linguistics compared to other aspects of the data). But,

since many were and are proficient in Sanskrit, Indian intellectuals could follow the philological arguments produced in support of the Indo-Aryan migrations, and their analysis of this aspect of the data in the 19^{th} century soon contributed to what was to become and remain a wide-scale rejection of the entire hypothesis. While opposition to the theory is nowadays primarily fueled by archaeologists, the first voices of suspicion were raised by Sanskrit scholars.

3. Malleable evidence and circular explanations of the data: some examples

3. 1. The by-now well-known snub-nosed *dāsya* case best exemplifies the malleability of this type of evidence and especially deepened the suspicions of a number of Indian scholars about the bias evidenced by western Indologists in their selection of certain interpretational possibilities to the exclusion of others. It involves the Sanskrit terms *anāsa*, an adjective used in a solitary instance in the *Rgveda* to describe the *dāsyas* as enemies of the *ārya*s. The term allows a rendering of *a-nāsa* 'without nose', or of *an-āsa*, 'without mouth,' or 'uncouth', as Sayana, the earliest commentator on the *Rgveda* had construed it. But, in the racially-surcharged context of nineteenth century philology, with certain European scholars combing through the Vedic texts in search of racial clues of an Aryan invasion, the former option of nose-less was appropriated and standardized as evidence of an aboriginal pre-Aryan nasal type. The ensuing inference, that this reference was to a snub-nosed aboriginal type in contrast to a distinct straight-nosed Aryan and therefore non-indigenous type, seemed natural. We must recall that the equation of language with race was rarely brought into question until relatively recently. Again, a quote best illustrates how this type of deductive reasoning was received in certain circles (Iyengar 1914: 6)

> One solitary word *anāsa*, to the *Dāsyu* has been quoted by...Max Müller...among numerous writers, to prove the *Dāsyus* were a flat

> nosed people, and that, therefore, by contrast, the Aryas were straight-nosed. Indian commentators have explained this word to mean *an-āsa*, devoid of fair speech....[T]o hang such a weight of inference as the invasion and conquest of India by the straight nosed Aryans on the solitary word *anāsa* does certainly seem not a very reasonable procedure

The circularity here (and in other similar interpretations of the Vedic texts) should be clear: reading the texts with the presupposition that IE peoples had invaded India, a highly questionable (and now considered untenable) racial reading was derived from a few innocuous words, which were then proffered in support of the initial theory. Indeed, the more one grapples with the data, the more one comes to realize that much of the evidence associated with the Indo-Aryans, whether philological, linguistic or archaeological, is extremely ambiguous and can thus be construed in sometimes diametrically opposed ways by scholars approaching the issue with differing intellectual backgrounds, and cultural orientations.

Linguistic paleontology, for example, is another aspect of the linguistic evidence that has probably been the most conspicuous and durable in the history of the homeland quest, even as few linguists today put much stock in the method. Linguistic paleontology involves hypothesizing about the social, religious, political, economic, ecological, cultural and geographical environment of pre- and proto-historic cultures from linguistic fossils, or cognate terms, preserved in the various members of a language family. While the premise of the method is actually quite brilliant, there are simply far too many problems involved in relying on it for determining the specifics of any supposed homeland. From well before the time of Keith, who, in 1933, cautioned that: "it should be a warning when we find that linguistically we may assert that the Indo-Europeans knew butter but were unacquainted with milk,...knew of snow but not rain, and feet but not hands," the method has been subject to a well-known litany of complaints by linguists over the decades (see for example Fraser 1926; Dhar 1930; Pulgram 1958; Colemann 1988; Anthony

1995; Krell 1998; Marcantonio 2002/5, etc). One way to realize the circularity inherent in much of this debate is to consider how such data can be construed to support either a migration from West to East or from East to West, with little or nothing in the data itself to favor either view. Marshalling the data in support of one of these positions, then, is a consequence of a starting premise, rather than a consequence of the data itself. But, typically, the selected reading of the data is then configured to support the initial presupposition. Examples from linguistic paleontology abound: for example, the common Migrationist argument that exotica unique to the East, such as camels and elephants, do not have cognates in other IE languages. They are therefore not reconstructed to PIE. The PIE homeland could thus not have been in the places where such exotica are to be found, namely, in the East. But as has been pointed out, especially by Indian scholars, since the time of Max Müller (1887: 101), who favored a Bactrian homeland: "suppose that the elephant and the camel had really been known ... by the united Aryans, when living in Asia, would it not have been natural that, when transplanted to the northern regions, their children who had never seen a camel or elephant should have lost the name of them" – in other words, what if the terms of Eastern exotica had been lost by all the other IE languages upon leaving an Eastern homeland? As an aside, and ironically, the Russian linguist Gamkrelidze & Ivanov (1983b) have more recently construed elephants, camels, and even monkeys to IE although this view is not without its detractors. The same problem holds good for situating the homeland in the NW on the grounds of the oft-cited salmon and birch tree evidence common to the Western languages – what if these terms had been coined *after* leaving an Eastern homeland?

3. 2. The same ambiguity applies to the Indo-Aryan terms sprinkled all over Central Asia and the Near East. There are Indo-Aryan loans In the Finno-Ugric languages, in the Mitanni and other documents unearthed In the Near East from the 15th century BCE, and in place and river names in Central Asia. Here, too,

there is little preventing cynics from re-construing the evidence in a manner completely opposed to the dominant 'Migrationist' model of interpretation. In other words, these linguistic clues could either have been left by Indo-Aryans migrating East from a Western homeland, or emigrating West from an Eastern homeland with very little in the data itself to favor either trajectory. The point of all this is not that any of the above arguments in any way support an emigration of the Indo-Europeans from an Indian homeland to the West, a proposal which I feel has usually provided little of positive with which to recommend itself (although, see Kazanas in this volume). The point is simply to underscore how the initial interpretative model both frames the reading of the data, and is itself framed by it in circular fashion. In any event, linguistic paleontology is certainly not the type of data that is likely to convince those suspicious of the ambiguity of much of the linguist evidence that has typically been marshaled to insist on the migration of the Indo-Aryans into the subcontinent from an external homeland.

The 'substratum' evidence in Vedic texts is another linguistic example of how data relevant to the early history of the Indo-Aryans is read through one interpretative model to the almost complete exclusion of others, a reading that is then used as a cornerstone for the initial model itself. A substratum in this case refers to an indigenous language being subsumed and displaced by an alien intrusive language. In this process, the indigenous language affects the dominant language by depositing into it its own linguistic features, such as vocabulary or morphology. These linguistic features form a substratum in the dominant language, which can be discerned by diligent linguists even after many centuries. In the Vedic case, there are non-IE linguistic features in the Vedic texts suggesting a language community preexisting that of the Vedic speakers who must therefore have been intrusive from elsewhere. In addition to vocabulary, the most common non-IE features referred to in this regard are: phonologically, the retroflexes; morphologically, gerunds; and syntactically, a post-posed 'quotative' marker, *iti*. However, apart from the fact that

nearly all such linguistic features have at one time or another been argued to represent spontaneous *internal* developments (Hock, 1975, 1984a, 1996) rather than borrowings, there has been very little discussion as to whether such influences between these languages – whether they be lexical or syntactical – if accepted as borrowings, can *only* be the result of substratum. What of the possibility of ad-stratum, or even super-stratum relationships perhaps adequately accounting for them? An ad-stratum situation is where languages influence each other due to being geographically adjacent to each other, as opposed to one being superimposed on another. This is not mere semantics: a substratum model points to an immigration of Indo-Aryans, an ad-stratum model is amenable to an Indigenist position. After all, and importantly, in the verifiable historical period, the Dravidian languages in the South of India and the Indo-Aryan ones in the North are continuing to greatly influence each other linguistically, but this has nothing to do with substratum – ad-strata or various expressions of bilingualism are the principal factors involved. Since ad-stratum accounts for the on-going influences between Dravidian, Munda and the Indo-Aryan languages in the present day, how do we exclude the possibility that a similar situation might not also account for the shared influences between these languages in the less-verifiable pre- and proto-historic period?

Now, since the Dravidians speakers are also generally considered immigrants into the subcontinent, the suggestion, first made by Bloch (1924, 1928) that perhaps these relevant linguistic features could have been borrowed by Indo-Aryans from invading Dravidians, rather than by Dravidians from invading Indo-Aryans is another possibility that has yet to receive scholarly attention. One significant exception, and of significance here, is Witzel's (1989, 1999, 2000, 2005) conviction that there was no Dravidian influence in the early *Rgveda*. Thus the subsequent influences of Dravidian on later Indo-Aryan were the result of an ad-stratum relationship between these two languages. If the Dravidian influences in the Old Indo-Aryan Vedic texts could have been caused by ad-stratum (or super-stratum) and this process has

continued ever since with the New-Indo-Aryan languages and the Dravidian ones, then how do we exclude the possibility that all the various innovations visible in Indo-Aryan, whether Dravidian, Munda, Masica's language *x* (1979) or anything else, could not have been the result of these other languages intruding onto an Indo-Aryan speaking area as super-stratum as opposed to vice-versa, or of acting as ad-stratum as in the modern linguistic area situation of the subcontinent? As first argued by Bloch (1924, 1928) the Northern pocket of Brahui-speakers could easily represent a linguistic community moving from the South to the North (or the remnant of an intruding Dravidian speaking community, which need have no bearing on the origins of the Indo-Aryans). In short, convergence, or any type of borrowing or similarities between languages, whether lexical or structural, in and of itself does not *necessitate* a situation of a linguistic substratum. That Indo-Aryan intruded onto a non-Indo-Aryan substratum still has much to recommend itself. Perhaps it is the least complicated way of accounting for the available evidence. But it is nonetheless based on the acceptance of certain assumptions that can be called into question. The evidence is malleable.

As for the loan words (i.e. not derivable from normative IE roots) in the Vedic texts pertaining to flora and fauna (Masica, 1979; Kuiper, 1991), or even personages (Kuiper, 1991; Witzel, 1995b, 2000), while these may very well bear evidence of an indigenous linguistic substratum pre-existing a migration of Indo-Aryan speakers, what of the possibility that some of these non-IE clues may rather represent exotic items introduced to the subcontinent by trade, and that the names might simply not reveal the presence of non-Indo Aryan immigrants? In other words, how does philology help us eliminate the possibility that it is these items and personages who are intrusive into the subcontinent, rather than the Indo-Aryans? I have discussed all these substratum issues in far more detail elsewhere (1991, 2001) and do not wish to duplicate the same material, but I refer to them again here simply to underscore the danger of circularity: that is, of a monolithic

interpretative model consuming all the data, which is then marshaled in support of the original model.

3. 3. While some of the above examples have hopefully given some sense as to how some of the philological and linguistic evidence can be reconfigured when viewed by those disposed to challenge established paradigms, that is, how easily the direction of circularity can be flipped around, most of the arguments raised by the detractors of the Indo-Aryan Migration theory deal with the archaeological evidence. Partly due to the fact that there has been so little opportunity available for the study of historical linguistics in India, in contrast to the fact that the Archeological Survey of India is a well-established scholarly entity, it has been archaeologists who have been the most vociferous in their opposition to the theory. Their position is fairly straightforward: there has been no success at finding any conclusive tangible archaeological evidence of an intrusive cultural presence into the subcontinent corresponding to that of the Indo-Aryans in anything approaching the time period required by the normative theory. (Tellingly, Indian archaeologists who do still support the migration hypothesis typically tend to defer to the 'linguistic evidence'). Since the Aryans have not been identified through positive evidence, the issue in archaeological circles has consequently revolved around whether the presence of Aryans can at least be excluded by negative evidence, in other words, on the grounds of a lack of Indo-Aryan related evidence such as horses or chariots in certain geographical and temporal archaeological contexts. The main debate focuses on the cultural and linguistic identity of the Indus Valley Civilization for obvious reasons, given its time bracket of circa 3300 BCE - 1900 BCE immediately prior to the date of the supposed Indo-Aryan migrations.

One of the pillars of the Migrationist argument denying the Indo-Aryans a presence in the Indus Valley Civilization is the fact that the early Vedic texts seem to make no explicit mention of urban centers. How, then, could the Indo-Aryans have been the progenitors of such a sophisticated civilization? This seems a

fairly reasonable argument, but, on closer analysis, there are a few important points that merit consideration. The first pertains to the interminable discussions involving the Vedic horse: since the site reports of Sir John Marshal (1931), the pioneer of archaeology in India, the absence of horse remains in Indus sites has been the mainstay of the belief that the speakers of the Vedic language must have succeeded the Harappan civilization. The horse is clearly an animal highly-valued in the Vedic world, and thus it is perfectly reasonable to expect that if the Aryans were native to the Indus Valley, their presence should be evidenced by remains of the horse there. This lacuna in the Indian archaeological record tends to haunt any efforts attempting to correlate the Indus Valley civilization with the Vedic culture, which is a horse-using one; and, of course, any attempt to argue for an Indian *Urheimat*. Thus, one cannot fault scholars using the first appearance of horse bones in the archaeological record as an approximate *terminus post quem* for the Vedic texts. Since this animal has become almost synonymous with the Vedic Aryans and, by extension, the whole Indo-Aryan migration debate, the horse is presently "the most sought after animal in Indian archaeology" (Sharma 1974: 75).

The exact species of the equid is the crucial issue in these identifications. A number of claims of horse bones have been made over the years, at least one of which was accepted by a horse specialist, Bökönyi (1997; see also Sharma, 1974). But other experts, such as Meadow (1997), remain unconvinced, since the remains could have appertained to either *Equus caballus* Linn, or to some other member of the horse family. As an aside, to pronounce unambiguously that there is *no* evidence of the horse in the IVC, on one side, or to insist that there is conclusive evidence, on the other, are both somewhat sleights of hands. A more accurate statement would be that there is some evidence, but it has been contested. There is no *uncontested* evidence because there are only minor differentiating features between the various species of *Equus*. *Equus hemionus khur*, for example, is indigenous to the Northwest of the subcontinent, but it is *Equus caballus* that is the sought-after Aryan steed. The differences between the species are

either difficult for experts to identify or, unless the specific distinguishing parts of the skeleton arc found (certain teeth and the phalanges, toe bones) impossible to determine with certainty. Obviously, the horse would never have been an issue had it not been linked with the Aryans; were it not for the politicization of the issue, the reports of horse evidence in the Indus Valley, albeit sparse, would hardly have raised any eyebrows. Nevertheless, unless horse bones are undeniably found in the early, pre- and Mature Harappan strata, the Indo-Aryan speakers may be (and already are) allowed a degree of synthesis with the later Harappan civilization, but their status as intruders, albeit perhaps earlier than previously held, will still not be considered convincingly undermined to the satisfaction of all.

But, irrespective of the differences between Meadow and Bökönyi, a caveat is in order, here: to say that the Indo-Europeans and the Indo-Aryans knew the horse and must therefore have come from a place where the horse is indigenous is an assumption that can be brought into question. An item can be central to a community and still be an import from somewhere else. The horse has always been highly valued in Indic texts; from the Vedic, through the Epic, and up to the Sultanate period, it has always been an elite item, and it has always been an import from the Northwest and never indigenous. One must accordingly be wary of making the Indo-Aryans themselves overly synonymous with the horse, since the horse could have been imported in the proto-historic period, just as it has been throughout the historic period, but this in itself need not a priori indicate that the Indo-Aryans were imports as well. The lacuna in the archaeological record could well be due to the elite and rare nature of this beast which may have been the preserve of martial and sacrificial contexts and thus unlikely to turn up in large quantities in normal archaeological contexts.

The chariot is a good example of the limitations of relying on the archaeological record for the purposes of dating a cultural group. The earliest archaeological evidence of the chariot is sometime between the fourth and first century BCE (Gaur 1983).

If we accept the dates assigned by most scholars to the *Rgveda*, viz., 1500-1200 BCE, the chariot as known to this text must have unquestionably been in existence on the subcontinent for a thousand years or so before becoming evidenced archaeologically just before the common era. The archaeological *argumentum ex silentio* clearly shows its limitations in this period during which we know the chariot was extant from the literary evidence but it has not been verified archaeologically. Obviously, the further back in time we go, the more the likelihood of finding such iconographic evidence decreases. Therefore the paucity or absence of an Indo-Aryan diagnostic item in the archaeological record might not mean the absence of that item and nor, therefore, of the Indo-Aryans, otherwise the chariot evidence would compel us to date the appearance of the Indo-Aryans on the subcontinent to after the 4th century B.C.E., which is patently absurd. The excavations at Mehrgarh are another example: they threw the date for evidence of agriculture back *two entire millennia,* and subverted the view that agriculture and urbanization were both diffused from West Asia. This clearly underscores the danger of establishing theories predicated on *argumentum ex silentio* in the archaeological record. Having said this, however, it seems fair to add that the burden of proof lies with anyone arguing for the existence of the horse in the Indian subcontinent in the 3rd and 4th millennium BCE – which an Indigenous case must argue.

The other argument that usually surfaces in denying the Indo-Aryans a presence in the Indus Valley Civilization, again since the time of Marshall, is the fact that the early Vedic texts seem to make no explicit mention of urban centers. However, even if the composers of the Vedic hymns did not primarily live in the cities of the Indus Valley, this fact in and of itself does not mean that the Harappans could not have been Indo-Aryan speakers. A language family can obviously encompass urban dwellers as well as village dwellers and nomadic pastoralists as is obvious all over the world today. Anyone promoting Dravidian, which has been by far the most common assumption made, or Munda, as the language of the Indus Valley implicitly accepts such a situation: namely,

urban Dravidian or Munda speakers co-existing with non-urban tribal Dravidian or Munda speakers, because the southern Dravidian-speaking culture and eastern Munda-speaking culture were radically different from that of the Indus Valley in the 3rd millennium BCE. So if hypothetical urbanized Harappan Dravidians or Mundas could have coexisted on the subcontinent with their non-urbanized tribal linguistic brethrens to the South or East, then Indo-Aryan speakers could have done likewise. Therefore, even if the early Vedic texts do not elaborately describe the flourishing cities of the Indus, it is not clear how one would discount the possibility that they could still have been the product of an Indo-Aryan pastoral society that co-existed with an Indo-Aryan urban one.

Secondly, there are two obvious routes to the Indian subcontinent from any external homeland: a northern route from the Northeast of the Caspian Sea through the Steppes of Central Asia and down through Afghanistan and into India, and a southern route from the Southeast of the Caspian Sea through the deserts and plains of North Iran and into Afghanistan. (And one can immediately note that there is no more consensus regarding the identification of the Indo-Iranians in the archaeological record, than that of the Indo-Europeans or Indo-Aryans). Both the northern and southern routes have diametrically differing archaeological cultures, which have been identified with the Indo-Aryans in a variety of ways by different scholars (see for example Parpola & Carpelan (2005) for the Northern Route; Sarianidi (1987) for the Southern). The most obvious problem with subscribing to the Northern route is that no traces of Steppe cultures have been detected south of the Hindukush. The main obstacle facing scholars arguing for a southern route is that the archaeological culture of the Gorgan Plain in southern Turkmenistan and northeastern Iran is littered with the sites of sophisticated urban centers during the proposed time frame of the migrations. How does this fit with the assumed pastoral non-urban nature of the Indo-Aryans? In either event, whether one favors the northern route or a southern one, one will nonetheless have to

accept that the Indo-Aryans passed through the BMAC, 'Bactria Margiana Archaeological Complex', before entering the subcontinent. Some have argued that the Indo-Aryans were actually the founders of this civilization (Sarianidi, 1993 a & b). Given this, attention must be drawn to the fact that the BMAC was a sophisticated civilization consisting of fortified towns. Structures identified as temples, were 'monumental'. Indeed, according to later excavators of BMAC sites, the area appears to mirror a facsimile of state structured polity of power (Hiebert 1995, 1998; Hiebert & Lamberg-Karlovsky, 1992) Any acceptance that all this was the handiwork of Indo-Aryans, or that the Indo-Aryans passed through this area, which they must have done from *any* point of origin outside the subcontinent, during the assumed time frame of the migrations in the 2nd millennium BCE, will entail abandoning certain stereotypes such as that the Indo-Aryans knew no urban centers. This further underscores the point that lack of urban references in the *Rgveda* is not an acceptable indicator of the Indo-Aryans' ignorance of such things. Add to this that, in fact, Pirak, a site in the Punjab in the first half of the second millennium BCE – right when and where many Migrationists would place the intruding Indo-Aryans – is a town of some size and it is hard to avoid the conclusion that the hymns may not be sparse in urban references due to ignorance or unfamiliarity with urban centers. Moreover, one might also note that although the horse was certainly utilized in the BMAC as attested by its representation on grave goods, no horse bones have been discovered there (despite the fact that an unusually high number of animal bones have been found). Therefore, absence of horse bones does not equal absence of horses, nor necessarily of Indo-Aryans.

4. Conclusion

In my view, it will only be the decipherment of the script that will prove decisive in this whole issue to the satisfaction of most scholars. Ultimately, the answer to the linguistic identity of the Indus Valley, which has decisive implications for placing the Indo-

Aryans, lies in our hands, but it has yet to yield its secret, despite a plethora of attempts. If the Indus script turns out to represent an Indo-Aryan language during this period there would be massive implications and corollaries for the entire IE homeland problem, especially since shards found suggest that the script may go back as far as 3500 BCE, and most Indo-Europeanists hold that the Indo-Europeans were still undivided till sometime between 4500 - 3500 BCE. In other words, an Indo-Aryan script on the subcontinent at a time frame when the proto-Indo-Europeans were supposed to still be more or less undivided undermines almost all homeland proposals with their ensuing migrational trajectories of the IE speakers. And all the ink spilt on attempting to date the Veda to 1200 or 1500 BCE will merit the skepticism that Indigenists have generally directed to such efforts. Actually, as an aside, although the topic is beyond the scope of this paper, a controversy remains ongoing after over a century as to whether the Vedic texts contain astronomical clues pointing to a date at least a millennium or millennium and a half earlier than the date found in text books. All in all, two centuries worth of IE homeland speculation will be subverted overnight if an Indo-Aryan script emerges – we might just note as an aside that Ventris was quite surprised to see Greek emerge from the linear B script from Crete, since, as with the Indus script, the prevailing assumption at the time was that it must contain a pre-IE language.

On the other hand, if the script turns out to be any language other than Indo-Aryan, then the Indigenist position no longer merits further serious scholarly attention. In my opinion, this eventuality will be the only development that will convince a large number of Indian scholars that the Aryans were, indeed, immigrants into India. Whatever language is contained in the script, it would be unwise for decipherers, in my opinion, to eliminate either Sanskrit, Dravidian or Munda as possible candidates. The answer, after all is said and done, is written on the seals.

To conclude this cursory discussion of the data, then, the idea of an Aryan immigration into India remains a plausible and in

places sometimes compelling way of accounting for at least some of the presently-available evidence. By Occam's razor, it requires less strain in accounting for at least some of the data. However, the fact that an Indigenous case can be made at all, in my opinion, underscores the malleability and circularity of the evidence, and requires that the less ideologically driven voices from the Indigenist school cannot be denied representation in discussions concerning Indian proto-history. Indo-Europeanists will likely consider the Indigenist position a very myopic view since it restricts itself solely to the history of only one member of this huge language family, with almost no concern as to how the other members got to be where they are. Indeed, the most serious obstacle the Indigenist view faces is that, like it or not, it must inherit the corollary of arguing for an IE language group indigenous to India: if the Indo-Aryan speakers are autochthonous to South Asia, then how are they related to the other IE languages? The corollary, at least from the realistic perspective that monopolizes present interpretational paradigms on South Asian history, can only be that the other Indo-Europeans left the subcontinent for the West. But this position has so far produced very little with which to recommend itself. Even granted that there are significant problems found with all the IE homeland proposals, clearly, an Indian homeland theory is even more subject to the type of criticism that Indigenists have vented on other homeland theories, because very little serious evidence has been offered to support it. The Indigenist school tends to limit its concerns primarily to deconstructing the opinions of Migrationists rather than offering anything very positive that can account for the entirety of the IE problem. But this very myopia of the Indigenist position has nonetheless unearthed major problems with much of the evidence upon which the theory of Indo-Aryan migrations into South Asia has been predicated. This, in turn, however selective in scope, can nonetheless be seen as a contribution to the entire field of IE pre- and proto-history if it forces us to reevaluate some of our inherited assumptions, and prompts scholars to take a fresh look at the entirety of the evidence.

I leave it to Indo-European linguists to explore what the repercussions on this issue might be if criticisms of the methods of historical reconstruction such as those offered by Marcantonio in this volume begin to alter normative perspectives of the very value or validity of reconstructing a PIE, and the subsequent consequences this might have on the family tree/wave model it has spawned. But, as far as the Indic side of the family is concerned, all that can be determined with certainty in my view is that the Indo-Aryan languages were spoken in the North-West of the subcontinent at the dawn of proto-history as, of course, were other IE languages elsewhere. Despite the enormous amount of linguistic, philological and archaeological evidence that has been uncovered over the last two centuries, our ability to explain the source of these language connections has not generated anymore consensus than Max Muller's candid appraisal in 1838, that all that can be said about it is that it arose "somewhere In Asia and no more". Perhaps it is time for a new paradigm, or at least time to revisit basic presumptions on the matter. Archaeologists, philologists and linguists focused on South Asia have been unable to reach a consensus with regards to the origins of the Indo-Aryans. Indeed, the debate has become a source of exasperation to many due to the extraordinarily emotional and polemical tenor of the exchanges. The Indo-Aryan migration theory was and remains essentially predicated on the discoveries of Indo-European linguists, specifically, on the connection of Sanskrit with other IE languages. A fresh revisiting and reevaluation from Indo-European linguists of the basic presuppositions of the field would be a welcome and timely contribution, certainly to the field of Indian proto-history.

Endnotes

[1] Lal, B.B. "Let not the 19th Century Paradigms Continue to Haunt Us!" Inaugural address delivered at the International Conference on South Asian Archaeology, held at University of Bologna, Ravenna, Italy, July 2-6th 2007
[2] The simplest of web searches will immediately reveal the emotionality surrounding this issue; as a consequence, discussion on the topic was banned from the primary Indological list-serve.
[3] http://en.wikipedia.org/wiki/Californian_Hindu_textbook_controversy
[4] *Hindutva,* as defined by Savarkar (1989), essentially entails supposing that the Hindu majority in India has the best right to define the ethos of the modern Indian nation state, because the religious traditions it follows took root on Indian soil. This is in obvious contradistinction with other communities, such as the Muslims or Christians, who owe spiritual allegiance to traditions and symbols that have their origins on foreign soil.

References

Anthony, David W.
1995 Nazi and Eco-Feminist Pre-histories: Ideology and Empiricism in Indo-European Archaeology. In P. Kohl & C. Fawcett (eds), *Nationalism, Politics and the Practice of Archaeology.* Cambridge University Press. 82-96.

Bloch, J.
1924 Sanskrit et Dravidien. *Bulletin de la Société de Linguistique de Paris* 76: 1-21.
1928-30 Some Problems of Indo-Aryan Philology. *Bulletin of the School of Oriental Studies* 5: 719-756.
1946 Brahui et Tsigne. *Journal of the Royal Asiatic Society* 12: 199-201.

Bökönyi, Sándor
1997 Horse Remains from the Prehistoric Site of Surkotada, Kutch, Late 3rd Millennium B.C. *South Asian Archaeology* 13: 297-307.

Bryant, E.F.
2001 *The Quest of the Origins of Vedic Civilization: the Indo-Aryan Invasion Debate.* Oxford /New York: Oxford University Press.

Bryant, E. F. & Patton, Laurie
2005 *The Indo-Aryan Controversy: Evidence and Inference in Indian History.* London: Routledge.

Coleman, Robert
1988 Book review of *Archaeology and Language* by Colin Renfrew. *Current Anthropology* 29/3: 449-453.

Dhar, Lachhmi
1930 *The Home of the Aryans.* Delhi: University Publications.

Fraser, J.
1926 Linguistic Evidence and Archaeological and Ethnological Facts. *Proceedings of the British Academy* 12: 257-272.

Gamkrelidze Thomas V. & Ivanov, Vjaceslav, V.
1983a The Ancient Near East and the Indo-European Problem: Temporal and Territorial Characteristics of Proto-Indo-European Based on Linguistic and Historico-Cultural Data. *Soviet Studies in History* 22:1-2:7-52.
1983b The Migration of Tribes Speaking the Indo-European Dialects from their Original Homeland in the Near East to their Habitations in Eurasia. *Soviet Studies in History* 22: 53-95.
1985a The Problem of the Original Homeland of the Speakers of Indo-European Languages. *Journal of Indo-European Studies* 13: 174-184.
1985b Response to I. M. Diakonoff (1982). *Journal of Indo-European Studies* 13: 175-202.
1990a The Early History of Indo-European Languages. *Scientific American,* March, 110-116.
1990b On the Problem of an Asiatic Original Homeland of the Proto-Indo-Europeans. In T. Markey & J. Greppin (eds), *When Worlds Collide.* Ann Arbor: Karoma.
1995 *The Indo-European and the Indo-Europeans.* Trends in Linguistics Studies and Monographs 80. Berlin: de Gruyter.

Gaur, R.C.
1980-81 The Aryans: A Fresh Appraisal. *Puratattva* 12: 133-135.
1993 *Excavations at Atranjīkherā.* Delhi: Motilal.

Gimbutas, M.
1966 Proto-Indo-European Culture: The Kurgan Culture during the Fifth, Fourth, and Third Millennia B.C. In George Cardona, Henry Hoenigswald & Alfred Senn (eds), *Indo-European and Indo-Europeans.* University of Pennsylvania Press.
1974 Archaeologist's View of PIE in 1975. *Journal of Indo-European Studies.*2: 289-307.
1975 The First Wave of Eurasian Steppe Pastoralists into Copper Age Europe. *Journal of Indo-European Studies* 5:277-338.

1985 Primary and Secondary Homeland of the Indo Europeans. *Journal of Indo-European Studies* 13: 1& 2: 185-202.
1997 *The Kurgan Culture and the Indo-Europeanization of Europe.* Edited by Miriam Dexter & Karlene Jones-Bley. Washington D.C.: Institute for the Study of Man.

Golwalkar, M. S.
1947 *We, or Our Nation Defined.* Nagpur.

Hastie, W.
1882 *Hindu Idolatry and English Enlightenment.* Calcutta: Thacker.

Hiebert, Fredrik T.
1995 South Asian from a Central Asian Perspective. In Erdösy, G (ed.), *The Indo-Aryans of Ancient South Asia.* New York: de Gruyter. 192-205.
1998 Central Asians on the Iranian Plateau: A Model for Indo-Iranian Expansion. In: Mair, V. H. (ed.) *The Bronze Age and Early Iron Age Peoples of Eastern and Central Asia* I. Washington D.C.: Institute for the Study of Man. 148-161.

Hiebert, Fredrik T. & Lamberg-Karlovsky, C.C.
1992 Central Asia and the Indo-Iranian Borderlands. *Iran.* XXX: 1-15.

Hock, Hans Henrich
1975 Substratum Influence on (Rig-Vedic) Sanskrit? *Studies in the Linguistic Sciences* 5/2: 76-125.
1979 Retroflexion Rules in Sanskrit. *South Asian Languages Analysis* 1: 47-62.
1982a The Sanskrit Quotative: A Historical and Comparative Study. *Studies in the Linguistic Sciences* 12 /2: 39-85.
1982b Aux-Clitization as a Motivation for Word Order Change. *Studies in the Linguistic Sciences* 12 /1: 91-101.
1984a (Pre-)Rigvedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. *Studies in the Linguistic Sciences* 14/1: 89-107.
1984b *Principles of Historical Linguistics.* Berlin: de Gruyter.
1993 Subversion or Convergence? The Issue of Pre-Vedic Retroflexion Re-examined. *Studies in the Linguistic Sciences* 23/2: 73-115.
1996 Pre-Rgvedic Convergence between Indo-Aryan and Dravidian? A Survey of the Issues and Controversies. *Ideology and Status of Sanskrit.* Edited by Jan E. M. Houben. Leiden: Brill.
1999a Through a Glass Darkly: Modern "racial" interpretations vs. textual and general prehistoric evidence on *Ārya* and *dāsa* /*dāsyu* in Vedic Indo-Aryan society. *Aryan and Non-Aryan in South Asia.* Harvard Oriental Series; Opera Minora 3: 145-174.

1999b Out of India? The Linguistic Evidence." *Aryan and Non-Aryan in South Asia.* Harvard Oriental Series; Opera Minora 3: 1-18.
2005 Philology and the historical interpretation of the Vedic texts. In E. Bryant & L. Patton (eds). 282-308.

Iyengar, S.
1914 Did the Dravidians of India Obtain their Culture from Aryan Immigrant [*sic*]. *Anthropos* 1-15.

Jones, W.
1788 On the Gods of Greece, Italy, and India. *Asiatic Researches.* 1: 221-275.

Keith, A. B.
1933 The home of the Indo-Europeans. In Pavry, J. D. C. (ed.), *Oriental Studies in Honour of Cursetji Erachji Pavry.* Oxford University Press. 189-199.

Krell, Kathrin, S.
1998 'Gimbutas' Kurgan-PIE Homeland hypothesis: a Linguistic Critique. In Blench, R & Spriggs, M (eds), *Archaeology and Language II.* London: Routledge. 267-289.

Kuiper, F.B.J.
1955 Rigvedic Loan Words. *Studia Indologica.* Bonn. 137-185.
1967 The Genesis of a Linguistic Area. *Indo-Iranian Journal* 10: 81-102. Reprinted in *International Journal of Dravidian Linguistics* 3/1: 133-153.
1991 *Aryans in the Rigveda.* Leiden Studies in Indo-European 1. Atlanta: Rodopi.
1995 On a Hunt for Possible Objections. *Indo-Iranian Journal* 38: 239-247.

Legge, F.
1902 The Home of the Aryans. *Academy* 63: 710-711. Maine, Sir Henry Summer
1875 *Lectures on the Early History of Institutions.* New York: Henry Holt.

Marcantonio, A.
2002/5 Linguistic paleontology: Science or fiction? A case study within Uralic. In B. Winsa (ed.), *Finno-Ugric Peoples in the Nordic Countries.* Stockholm: Academia Tornedaliensis. 77-95.

Marshall, Sir John
1931 *Mohenjo-Daro and the Indus Civilization* 1. London: Arthur Probsthain.

Masica, C.
1976 *Defining a Linguistic Area.* Chicago: Chicago University Press.
1979 Aryan and Non-Aryan Elements in North Indian Agriculture. In Deshpande, M. & P. Hook (eds), *Aryan and Non-Aryan in India.* Ann Arbor: University of Michigan. 55-151.

Meadow, Richard
1997 A Comment on 'Horse Remains from Surkotada' by Sándor Bökönyi. *South Asian Archaeology* 13:308-315.

Müller, F. Max
1883 *India: What Can it Teach us?* London: Longmans.
1887 *Biographies of Words and the Home of the Aryas.* New Delhi: Gayatri.

Nichols, Johanna
1997 The Epicenter of the Indo-European Linguistic Spread. In Blench, R & Spriggs, M (eds), *Archaeology and Language I.* London: Routledge. 122-148.
1998 Eurasian Spread Zone and the Indo-European Dispersal. In Blench, R & Spriggs, M (eds), *Archaeology and Language II.* London: Routledge. 220-266.

Parpola, Asko & Carpelan, Christian
2005 The Cultural Counterparts to Proto-Indo-European, Proto-Uralic and Proto-Aryan: Matching the Dispersal and Contact Patterns in the linguistic and Archaeological Records. In Bryant, E & Patton, L. (eds). 107-141.

Prakash, Buddha
1996 *Rgveda and the Indus Valley Civilization.* Hoshiapur: Vishveshvarananda Institute.

Pulgram, E.
1958 *The Tongues of Italy.* Cambridge (Mass.): Harvard University Press.

Sarianidi, V.
1977 New Finds in Bactria and the Indo-Iranian Connections. In Taddei, M (ed.), *South Asian Archaeology.* Naples: Istituto Universitario Orientale. 643-659.
1987 South-West Asia: Migrations, the Aryans and Zoroastrians *International Association for the Study of the Cultures of Central Asia Information Bulletin* 13: 44-56.
1990 Togolok 21, an Indo-Iranian Temple in the Karakum. *Bulletin of the Asia Institute* 4: 150-156.
1993a Margiana in the Ancient Orient. *International Association for the Study of the Cultures of Central Asia Information Bulletin* 19: 5-28.

1993b Recent Archaeological Discoveries and the Aryan Problem. In Gail, A & Mevissen, G (eds), *South Asian Archaeology 1991.* Stuttgard: F. Steiner. 251-264.
1994 Margiana and the Indo-Iranian World. In Parpola, A & Koskikallio, P (eds), *South Asian Archaeology.* Helsinki: Suomalainen Tiedeakatemia. 667-680.

Savarkar, V.D.
1989 *Hindutva* [6th edition]. New Delhi: Bharti Sahitya. Sharma A. K.
1974 Evidence of Horse from the Harappan Settlement at Surkotada. *Purātattva* 7: 75-76.
1992-3 The Harappan Horse was buried under the dunes of ...? *Purātattva* 23: 30-34.

Witzel, Michael
1989 Tracing the Vedic Dialects. In Caillat, C (ed.), *Dialectes dans les Littérateurs Indo-Aryennes.* Paris: Institut de Civilisation Indienne. 97-265.
1995a Early Indian History: Linguistic and Textual Parameters. In Erdösy, G (ed.), *The Indo-Aryans of South Asia.* Berlin: de Gruyter. 85-125.
1995b Rgvedic History: Poets, Chieftains and Polities. In Erdösy, G (ed.), *The Indo-Aryans of South Asia.* Berlin: de Gruyter. 307-352.
1999a Aryan and non-Aryan Names in Vedic India: Data for the Linguistic Situation, c. 1900-1500 B.C. In Bronkhurst, J & Deshpande, M (eds), *Aryan and non-Aryan in South Asia.* Harvard Oriental Series 3.
1999b Aryan and non-Aryan names. In Bronkhurst, J & Deshpande, M (eds), *Aryans and Non-Aryans, Evidence, Interpretation and Ideology.* Harvard Oriental Series, Opera Minora 3: 337-404.
2000 The home of the Aryans. In Hintze A. & Tichy E. (eds), *Anusantatyai. Fs. Für Johanna Narten zum 70. Geburtstag.* Münchener Studien zur Sprachwissenschaft, Beihefte NF 19. 283-338.
2005 Indo-centrism: Autochthonous Visions of Ancient India. In Bryant, E & Patton, L. (eds), 142-178.

Verbal Inflection from "Proto-Indo-European" to the Indo-European Languages: A Matter of Coherence?

Paolo Di Giovine
University of Rome 'La Sapienza'

Everybody who has investigated the verbal inflection of the best known ancient Indo-European languages is aware that there is a coherent bundle of features which join together Old Indian, Avestan, Ancient Greek, Gothic, Old Church Slavic, Latin and Hittite. This is a necessary condition to reconstruct a Proto-Indo-European verbal structure. But such a procedure, although in fact necessary in order to trace back a common stage, underestimates quite a relevant number of both inflectional and functional categories which show divergences sometimes going back to doubtless different archetypes. The most striking differences divide the Hittite verbal inflection from the Indo-Iranian one, but there is also a set of features which mark Indo-Iranian and Greek in contrast to Latin and Germanic. I shall review these issues and attempt to show to what extent a difference springs up as an independent innovation or points instead to different archetypes. The latter case would suggest that even if 'a' Proto-Indo-European really existed, it was not a compact language, but at least a group of quite differentiated dialects.

1. How and to what extent the choice of the model of interpretation can affect the reconstruction of the Indo-European verbal system.

It has recently become evident to the careful reader that the numerous, initial models of interpretation adopted for the purpose of reconstruction within IE, the different "universi del discorso" ('reference points'), to use Coseriu's words (1997: 129 ff.), may actually be cause of concern. In fact, in order to implement a reconstruction, some scholars prefer to take into account a comparative corpus as wide as possible, drawn from all the IE

languages – therefore, independent of the more or less conservative nature of the various languages / branches. Other scholars instead consider more appropriate to use as a starting point the 'core' languages, that is, those languages within the family that display clearer and closer correlations. Obviously, at times drastically different conclusions can be drawn in choosing the one or the other initial model of interpretation, since a different measurement of weight is in this way attached to the relevant data. This is mostly evident at the morphological level, specifically in the field of verbal reconstruction. In order to avoid disorientation and perplexity when faced with such a variety of initial models, and related results, it is worth revisiting the whole issue, which can be summarized in the two points below:

1. What languages should be selected and taken into account for the purpose of reconstructing the IE verbal system, and how do we justify a given choice?
2. What conclusions can be drawn from a given, more or less fully implemented reconstruction of a prehistoric verbal system for the more general purpose of evaluating the unitary character, or otherwise, of that pre-historic entity that is commonly called: 'reconstructed Indo-European'?

I have discussed elsewhere the issue referred to in the first point (see Di Giovine 1977b); therefore in this article I shall only briefly mention (and up-date) the claims I made on that occasion. As to the issue mentioned in the second point, the concrete practice of reconstruction will give me the opportunity to revisit not only the genetic relationships within the family, but also the very model adopted to assess these relationships 'vertically', that is, the family tree diagram perspective[1].

2. The languages selected for the purpose of comparison and their weight within the reconstructed Indo-European verbal system.

It is a well known fact within the field that the paradigm of the IE verbal system, as first reconstructed by Bopp and then refined and codified in Brugmann's *Grundriss* (1913), has recently been called into question in view of the following factors: a) the discovery of new IE languages, mainly Hittite, but also Tocharian[2]; b) a deeper attention paid to those branches initially considered to be less important within the family, such as, mainly, Celtic. As a testimony of the challenges faced by the traditional paradigm one can mention the 'Indo-Hittite' theory, as proposed by Sturtevant in the 1920's (1929: 25 ff.), and subsequently taken up again by several scholars, among whom Cowgill (especially Cowgill 1974). According to this theory, the self-evident peculiarities of the verbal morphology of the Anatolian languages can be accounted for by postulating an early split into two major branches: the Anatolian branch and the branch consisting of all the other IE languages (those on the basis of which the neo-grammarian model has been formulated).

The thesis of an early split of the IE family can certainly constitute a plausible interpretation of the data, and, in fact, I shall deal with it in the second part of this article. On the other hand, whether or not the Anatolian branch plays a pivotal role in the reconstruction of IE is a question that has to be answered within the framework outlined in point (1) above, that is: what languages should be selected for the purpose of the reconstruction of the IE verbal system and what is their relevance, their 'specific gravity' – so to say – in this process. This question is even more relevant since several scholars (following the footsteps of Sturtevant), despite their reservations towards the Indo-Hittite theory as such, have started to call into question the reliability of a reconstruction of the verbal system that is based essentially on the correlations existing between Greek, Indo-Aryan and Iranian, and on the

decreasingly important role of the correlations existing among the remaining language groups (Latin and Italic, Germanic, Baltic and Slavic, Celtic, not to mention the 'minor' languages, such as Armenian and Albanian). Thus, these scholars, having observed that the simple inflectional system of the Anatolian verb corresponds to a reduced number of inflectional categories in the other IE languages – with the exception of Ancient Greek, Vedic and Avestan – have concluded that the Greek-Indo-Iranian complex verbal system must represent an innovation with regard to the oldest, original system. As representative of this new line of thoughts one could quote Meid (1975: 206-214; 1979: 161-170), Neu (1967: 221-230; 1976: 242-246) and Polomé (1985: 672-77 and 682-83), just to mention a few names[3].

It is clear that this issue has a relevance to the process of reconstruction, particularly when we are dealing with a reconstruction that is supposed to reach far back in time. Neither could one really make recourse to the thesis of a 'chronological stratification' within a reconstructed language, this clearly being just a 'rescuing procedure', with the help of which one can fit in all the relevant evidence and counter evidence. I have already pointed out elsewhere (Di Giovine 1996: 53-55) the methodological weaknesses embedded in a reconstruction that is in turn based on other reconstructions[4]. In fact, in this case, the reliability of the attained results decreases exponentially – one could imagine what an unlikely archaic Latin might have been reconstructed if nothing but the Romance languages were available. Not to mention that the pace of change of a given language in its pre-historical phase cannot be considered as constant, as is well known to those scholars who have tried to apply – definitively unsuccessfully – Swadesh's methods of lexical Glottochronology[5]. It is therefore understandable that such an extremely intricate issue – the (re-)evaluation of those linguistic sources at first considered to be marginal for the purpose of morphological reconstruction – has been mainly ignored in modern manuals of IE linguistics, including the (otherwise)

excellent manuals by Meillet (1937), Szemerényi (1990, where he attempts a classification into various chronological stratifications, in this way however expounding the topic in a much more complex way), up to Beekes (1995)[6]. Given that, in principle, each single IE language can contribute in equal measure to the general process of reconstruction, as well as to the reconstruction of the verbal morphology, priority in the choice of the linguistic evidence to utilize for this purpose should be dictated not by personal or practical motivations (such as: reference to the traditional bulk of studies, or easier to use languages), but by well founded criteria, as listed below:

a) The chronology of the linguistic data (obviously, the oldest records are more likely to preserve a linguistic picture closer to the pre-historical stage).
b) The extension of the records (that obviously reduces the risk of having to face a poor attestation of the verbal forms).
c) The structural situation of each of the languages involved, that is, in the case of the verb, the internal transparency of the forms and their easiness of interpretation.

The first two criteria are widely accepted and there is no point in dealing with them on this occasion. I shall therefore focus my attention on the analysis and discussion of the third criterion, to which little attention appears to have been paid (at least in open and explicit terms) by those scholars investigating the verbal morphology.

3. The typological structure of the Indo-European verb.

The structure of the verbal form – at the oldest stages of development of the various IE languages – is considered to be consisting, typically, of a root expressing the lexical value, and a series of affixes expressing morphological functions, the last of

which in the series is the inflectional ending (encoding person and number, as well as, at times, temporal and /or modal notions). This analysis is widely accepted, as testified by the now consolidated convention according to which it is the verbal roots that are listed as lexical entries in the IE etymological dictionaries. Similarly, it is reasonable and widely accepted to further classify the affixes (with the exclusion of the inflectional endings) into suffixes and prefixes (which are external to the root) and infixes (which are, instead, internal to the root[7]). The morphemes that are internal or represent an expansion of the root certainly belong to the oldest retrievable layer of IE, as is demonstrated by the fact that these morphemes are present not only in the 'Greek-Indo-Iranian core', but also in Germanic (*Ablaut* and reduplication, although somewhat blurred[8]), in Celtic, Latin and the Italic languages, and, to a greater or lesser extent, in all the other IE branches[9]. As suggested by Belardi (1990: 161), the word at the proto-IE stage must have consisted of a sequence of morphs, each one providing its own contribution to the overall meaning of the verbal (or nominal) form. Thus, looking at the verb from a typological perspective, we find numerous clues that allow us to postulate an oldest agglutinative / inflectional stage: agglutinative because we have a (not always linear) sequence of morphological markers; inflectional because the single morphemes may encode several functions at the same time, both 'cumulative' functions (for example in the inflectional endings) and mutually exclusive functions (for example in the *Aktionsart* suffixes[10]). Within this framework – a framework that still adopts the traditional, comfortable Humboldtian classification – one may observe a progressive *drift* from a compact, 'endomorphic' system towards a system that is (almost completely) 'exomorphic', that is, a system where the morphological markers tend to occur outside the structure of the root[11]. Through the passing of time the verb tends to loose its structural transparency: its internal structure tends to become opaque and its forms tend to crystallize, at the point that some morphological markers end up being transferred outside the

verbal lexeme onto the auxiliary verb (as is the case of the Romance and Germanic languages, several Middle and New Iranian dialects and, to some extent, also the Slavic languages). Not to talk about those extreme cases, such as Modern English, where the personal endings have been substituted by the subject pronouns.

This being the case, why would some scholars refuse to attribute an original, or at least, a more archaic structure to those languages that better preserve the transparency, and, therefore, the 'identifiability' of the building morphs (and lexical root) within the verbal phrase? The traditional way of reasoning appears to me to be quite well founded, and justifies the fact that the reconstruction of the IE verbal system can indeed be better achieved using as a starting point Vedic, Old Avestan and Greek, that is, those languages that (in addition to being attested in ancient records) have better preserved the original, structural model of the IE verbal formation. Obviously, this claim has to be double checked. In other words, keeping in mind the chronological aspect and the available records, one should be able to demonstrate that those languages that have evolved (to a greater or lesser extent) further away from the original, complex verbal structure have actually undergone a process of simplification (also at the level of inflectional categories; see below), a process of innovation that took place according to the 'language-specific *drift*' of the languages involved. This is exactly what I shall try to do in the following paragraphs.

4. The modal system: shared features and innovations.

The modal system across the IE area appears to be rather inconsistent – at least at a first glance – if one takes into consideration all the various IE branches. The obvious question is whether such an inconsistency is to be traced back to originally different structures, different archetypes or, on the contrary, to a

common source, in which case processes of formal and functional syncretism and / or processes of (mainly formal, but also, partially functional) innovation are assumed to have taken place.

The most articulated modal system (apart from Albanian, for which see below) is found in Vedic and Avestan, the most conservative languages within their respective language groups (Indo-Aryan and Iranian). In these languages, in addition to indicative, conjunctive, optative and imperative, another modal formation is documented, called 'injunctive' – a modal formation in reality closely connected to the indicative. As shown by Hoffmann (1967), the 'secondary' inflectional endings and the absence of augment, features proper of the injunctive, have been the starting point for the formation of the present indicative on the one hand (with its 'primary' inflectional endings), and that of the imperfect indicative (if the lexeme is imperfective) or the aorist (if the lexeme is perfective) on the other hand, through the pre-fixation of the augment. Thus, the injunctive would be a relatively older modal formation; it would have survived the coming into being of the indicative at the expense of a functional drift toward modalities that are a corollary of the imperative (such as 'prohibition', etc.). And, in fact, traces of injunctive are to be found not only in Avestan, but also, probably, in Greek, Armenian, and Celtic, as argued for by Lazzeroni (1977: 17-26).

In Greek the injunctive is no longer productive, surviving only as a relic (as mentioned above), but the other four moods are perfectly functional (this being the case also in Tocharian, where there is evidence for a four-ways modal system, including a still recognizable conjunctive and optative; see Pinault [1989: 124-29]). In Latin and Celtic, that preserve the indicative and the conjunctive, there is a functional syncretism of the categories of conjunctive and optative in the conjunctive (which, however, formally represents an old optative). The opposite holds in the Germanic languages, where only the optative remains (although not for long time); this category in turn subsumes also a conjunctive function. The optative has been preserved in Baltic,

although it has changed its function (compare the so-called 'permissive' in Lithuanian, for which see Stang [1942: 243-45]). In the imperative too (an autonomous mood) several innovative forms can be found – in general, Old Prussian appears to be the most conservative language (see Dini 1997: 89 e 264). A more drastic reduction of the modal forms is found in the Slavic languages, where only some old forms of optative survive (in addition to the indicative), although they are re-interpreted as imperative (Stang 1942: 238-42). In contrast, in Albanian the modal system is amply developed, although here we are mostly dealing with innovations; and, in fact, the moods that reflect an older stage of language (both formally and functionally) are only the indicative and the imperative.

At this point a closer look to the modal system of the Anatolian languages is in order. As is known, Hittite (as well as Luvian and the other Anatolian languages) have only two modal forms – the indicative and the imperative – and this fact (among others) has lead many scholars to assume that this must have been the original state of the art. Thus, on the one hand we have a system consisting of only two moods (a system older or at least contemporary to the more complex one), on the other hand we have a totally different, much richer, perhaps innovative tradition, a four / five ways modal system – hence the idea that the Greek-Indo-Iranian system might well be an innovation. It goes without saying that such an interpretation is based on the wide spread bias according to which what is simpler must be necessarily older (or, as it would be defined within the 'pseudo-anthropologic' model, more 'elementary'[12]). As a matter of fact, analysing the relevant Anatolian data from a comparative perspective it turns out that Anatolian has undergone a rapid process of syncretism in its entire verbal inflection, rather than preserving an (illusory) IE binary modal system, as pointed out by Lazzeroni (1995: 229):

> "Le lingue anatoliche partecipano all'innovazione panindoeuropea che ha prodotto le desinenze principali

> (cf. itt. *esmi* [...]), ma non conosce i paradigmi modali. Ma se è vero che questa innovazione si è attuata in presenza dei paradigmi modali, ciò significa che la tradizione indoeuropea confluita nelle lingue anatoliche conosceva i paradigmi modali; cioè che, per questo aspetto, essa non è più arcaica di quella confluita nelle altre lingue".[13]

This point of view regarding the evidence internal to Anatolian is actually backed up by what has been observed above with regard to the state of the modal system within the other IE branches, that is: a) the original five-ways (later on four-ways) modal system is actually documented very widely – from the Greek to the Tocharian area; b) those branches that have a reduced (two or three-ways) modal system display residual traces of moods that are no longer productive in their functionality; this in turn means that a syncretism of forms, and, often, also of functions, must have taken place. Not to mention that – at the typological level – complex systems (such as the nominal case endings and the modal endings) tend to undergo simplification and re-arrangement, as clearly shown by the drastic simplification that affected the nominal inflection in the transition from Latin to the Romance languages, or from Anglo-Saxon to Middle English (where also the personal endings of the verbs have been affected, having been completely replaced by the obligatory use of the subject pronoun[14]).

5. The tense-aspect and *Aktionsart* system.

The categories of Aspect, *Aktionsart* and Tense can reasonably be considered as the starting point for evaluating the verbal forms in the old IE languages. The importance of these categories has masterly been pointed out by Hoffmann (1970: 29ff.), and several other scholars through various publications[15]. Since the nature of these categories is rather well known I am not going to dwell about

this topic here. What is of relevance instead is to give a plausible answer to the following questions:

a) Did all the IE languages have these categories (at least in their oldest, retrievable stages), in which case they can be classified as an original, common feature belonging to a hypothetical, unitary stage of IE?
b) How are these categories codified in the various languages, and how and when each of them expanded at the expense of the other?

There appears to be full consensus among scholars with regard to the antiquity of the *Aktionsart* category (intensive, iterative, inchoative, etc.). Indeed, the inflectional categories expressing *Aktionsart* notions are very well documented in Indo-Aryan (the so-called 'secondary conjugation'), in Avestan, in the Anatolian languages as well as in Greek (although here they have lost their morphological autonomy), and in Germanic (where old, 'actional' stems have converged into the first class of weak verbs). Similarly, there is evidence for *Aktionsart* stems in Latin and in several suffixed forms in Baltic and Slavic. As to Hittite, its richness in *Aktionsart* stems (for which see Dressler 1968) has convinced several scholars, starting from Adrados (1974: 819 ff.), to reconstruct an IE verbal system based indeed on the opposition of actional stems, as is the case of Anatolian (notice however Jasanoff's (2003: 215) concerns towards this issue). As pointed out elsewhere[16], although suffixation (a typical 'exomorphic' process) is the most notable feature of the *Aktionsart* stems (stems in **-ye/o-*, **sk'e/o-*, **-n(e)w-*, etc.), these stems are also characterized by the root vocalism **-o-* typical of the deverbal stems (an 'endomorphic process'), as well as by reduplication, at least in those languages that have preserved it – reduplication involves an expansion of the root and is therefore another (basically) 'endomorphic' process.

Thus, if it is true that from a typological perspective exomorphism points to more recent formations whilst endomorphism points to more conservative ones[17], then the structure of the actional stems in the old IE languages suggests that these formations, as well as the categories of *Aktionsart* they encode, are relatively archaic. This being the case, we should turn now our attention to the other two functional categories – Tense and Aspect – that is, to those categories whose distribution appears to be less even within IE.

The notion of Aspect in the old IE languages, whose meaning is beyond doubt (contrary to what happens when we refer to Aspect in modern IE and non-IE languages[18]), is expressed through the binary opposition: imperfective *vs* perfective. This opposition is in turn codified morphologically through the opposition: present system *vs* aorist system – a structure present in Vedic, Avestan and, even more widely, in Ancient Greek. However, such a system barely survives in Latin and in Celtic, undergoes restructuring in Slavic, and tends to disappear elsewhere. The apparent absence of aspectually marked stems in Hittite and in the other Anatolian languages has lead some scholars (such as Schmitt-Brandt (1987) and Strunk [1994]) to doubt the antiquity of this functional category, whose formation should be therefore projected forward onto a later proto-IE period (Strunk 1994: 430). Thus, here there are two alternative ways of accounting for the linguistic evidence (this issue being particularly relevant for the current discussion): one way is to deny the antiquity of the aspectual notions (as argued for by Strunk e Schmitt-Brandt), the other is to deny the unitary character of IE (following the footsteps of Sturtevant [1929]). Before investigating this issue any further, let us first discuss the category of Tense and its morphology.

Those very scholars that deny the antiquity of the Present *vs* Aorist opposition tend in contrast to backdate the presence of the category of Tense in the verbal inflection of IE. As observed elsewhere (Di Giovine 1997a: 320 ff.), the formal encoding of

Tense is implemented through morphemes that are not particularly old, such as the augment, as well as through an opposition between primary and secondary inflectional endings, an opposition that is itself somewhat later (as indicated by the added 'particle' *-*i* denoting the *hic et nunc*, which produces the indicative starting from the injunctive). Actually, the fact that the opposition of Tense is absent in the various moods – apart from the indicative – in Indo-Iranian and Ancient Greek (the languages that have best preserved a modal system), whilst the aspectual opposition remains productive, should represent evidence that Tense was not the most relevant notion, at least not in a very old group of languages. At this point, the following, important question arises: since the category of Tense appears to be quite central in the other IE branches, could this mean that there must have been a split in the family within the field of functional categories? The answer appears to be negative, if we analyse the matter deeper. In fact, in Latin the category of Tense becomes the most relevant feature just at the time when the category of Aspect looses its prominence: the opposition *infectum : perfectum* – initially aspectual in nature (Present *vs* Aorist) – tends to become progressively 'temporal', as the *perfectum* subsumes within itself aspectual stems (Aorist) and *Aktionsart* stems (Perfect), both re-interpreted as tense stems. This is exactly what happens in Germanic, where the Preterit consists mostly of Perfect (actional) forms, and, to a lesser extent, of Aorist (aspectual) forms. Thus, in many languages one may observe the younger nature of Tense with respect to Aspect – this being also confirmed by the fact that the future tense comes into being later on, often having as a starting point the perfective stems. However, the picture is further complicated by the situation of the Anatolian languages. Here the absence of formal categories of Aspect is coupled with the fact that the temporal opposition in the verbal conjugation is expressed by the inflectional endings (which, by the way, cannot be derived the ones from the others). Nevertheless, I believe it is necessary to carry out a deeper investigation of the Hittite verbal system, to verify whether the picture presented

above – a verbal system poor in functional categories and strongly based on a temporal split between 'primary stems' and *Aktionsart* stems encoded by suffixes – is indeed a genuine picture. This issue has been raised on several occasions by Melchert (see Melchert (1997: 83-91) and Hoffner & Melchert [2002: 377-378]). Melchert has shown not only that several instances of lexical suppletivism are in fact of aspectual nature (as it also happens, not infrequently, in Greek and Indo-Aryan), but has also shown that several specific suffixes are markers encoding Imperfectivity rather than *Aktionsart* categories (compare in particular *-ye-/-ya-*, *-ške-/-ška-*, *-ešša-/-išša-*, *-anna-/-anni-*, with regard to Old Hittite[19]). If this analysis eventually proved correct (as is likely to be, given that it appears to be sound and is proposed by an authoritative scholar), we should be able to draw very different conclusions from the conventional ones whilst evaluating the evidence of the Anatolian (particularly the Hittite) verbal system for the purpose of IE reconstruction. In particular, one could state the following:

a) The Hittite verbal system, specifically at an older stage, did have three functional categories: Aspect, *Aktionsart* and Tense (obviously in addition to the categories of voice, person and number, all encoded through inflectional endings).

b) In this branch the category of Aspect has lost its relevance very early on, being no longer marked through inflection; as a result, only the Tense and the *Aktionsart* oppositions have been preserved.

With regard to point (a.), the correspondence between the Anatolian and the Greek-Indo-Iranian system is self-evident. It is important at this point to highlight that the category of Aspect is found in connection with verbal roots that are marked one way or the other, that is, roots that lacked the Imperfective or the Perfective, the opposition being therefore encoded at the lexical

level (a fact that, to my knowledge, has not yet attracted the necessary attention). This is a very old feature, a feature that must have originally characterized also the aspectual opposition of Greek and Vedic (partially also Avestan)[20], and whose traces are still retrievable in Latin, before the time when the high level of productivity of the derivative present stems and the sigmatic (or thematic) aorist stems triggered the expansion and full scale organization of the aspectual categories. In other words, according to this interpretation of the evidence, Hittite does not at all represent a language belonging to a lateral or parallel branch of IE. On the contrary, Hittite shows specific correspondences with those very languages traditionally considered to be more conservative in the area of the verbal system (as well as in other areas). Point (b.) too shows a good deal of convergence between Anatolian and the other IE languages: the category of Aspect looses its ground against an expansion of the *Aktionsart* stems on the one hand, and of the temporal opposition between Preterite *vs* Present on the other hand. This phenomenon – that is retrievable in the history of Hittite, the best documented Anatolian language – finds its perfect parallel in the typological development that takes place in most IE branches, from Latin to Germanic, to Indo-Aryan. Thus, we observe in the Anatolian languages that very same typological *drift* that has taken place elsewhere in a vast IE area, with the only difference that in Anatolian this *drift* was implemented much more rapidly. Put it in other words: the only noticeable difference between Hittite and the rest of the old IE languages is the fact that in Hittite the loss of Aspect took place before this opposition was properly implemented through the widespread introduction of perfective morphemes in imperfective roots or imperfective morphemes in perfective roots. This particularly rapid *drift* toward the loss of Aspect appears to be the only, real peculiarity of Hittite with respect to the other languages under discussion.

6. The other functional categories of the verb.

Before drawing conclusions from the data and analyses expounded above, and evaluating the impact these conclusions may have on the wider issue under discussion – the reconstruction of a more or less compact IE – it is necessary to briefly analyse the other functional categories of the verb in the IE languages. By this I mean the voice and the categories of person and number in the verbal conjugation, categories that are mainly encoded through inflectional endings in the old IE languages. In addition, one must take into account a split typical of the present tense: thematic *vs* athematic inflection, this split too being encoded through inflectional endings (although its functional meaning remains obscure).

Within IE the person and number categories are basically homogeneous (three persons and two numbers in all the old languages, with a third number, the dual, that is however in a receding phase already in the oldest stages of Indo-Iranian and Greek[21]). Furthermore, these categories are encoded in all the old languages through similar means, that is, through two series of inflectional endings – active and middle[-passive] –, so that any verbal form is clearly recognizable as being marked for a specific person, number and voice. There is a difference between the Greek-Indo-Iranian system on the one hand, and that of most of the remaining IE verbal systems on the other: in the first group middle endings are also used in close connection with the zero / reduced grade of the root (at least in the athematic stems). However, this divergence looses its relevance if one observes that this is in line with the *Ablaut* functionality of the root, a phenomenon that is in a residual phase in Anatolian and has been deeply restructured in the western IE languages (see Di Giovine 1995 and Carruba in this volume). The homogeneous character of the functional categories of person, number and voice across the old IE languages (including Hittite) is even more relevant in view of the 'typological revolution' that has taken place in the various, modern IE languages during a relatively short span of time, and whose most vivid testimony is offered by English. In this language

person and number are no longer encoded by inflectional endings, but by the personal pronoun subject, that is, an element external to the verbal form.

To complete the coverage of the topic, it is appropriate to take into account the opposition between thematic and the athematic conjugations, whose origin remains still unexplained. Also in this case, Hittite does not appear to distance itself from most of the other IE languages that have kept this formal opposition productive, or have at least preserved clear traces of it (such as Latin and the Old Germanic languages). As a matter of fact, within both its conjugations (the *-mi*-conjugation and the *-ḫi*-conjugation), Hittite is provided with thematic and athematic forms. As to the *-ḫi*-conjugation, all sort of arguments and counter arguments have been brought forward to account for it (for which see Jasanoff [2003]). If it is true that an indirect connection with the thematic conjugation of the other languages could possibly be established[22], it is also reasonable to envisage that the *-ḫi*-conjugation reflects an extremely complex situation, given the undeniable functional diversity of the verbs belonging to this class – this conjugation could be the result of a secondary convergence, or, even, a remnant of an hypothetical *$*$-h_2e* conjugation, as argued for by Jasanoff (2003: 217 ff.).

7. Is it appropriate to reconstruct a common Indo-European verbal system?

It is now time to draw more general conclusions regarding the genetic and mutual relations within the IE family, on the basis of the evidence offered by the IE verbal system (as put forward above).

First of all, the noticeable diversity of the functional categories observed across the IE languages at the synchronic, descriptive / typological level does appear in a different light when looked at in a diachronic perspective. The absence of the category of Aspect in many IE languages appears to be an innovative feature, rather than

an inherited one. The scarcity of moods typical of several 'western' languages on the one hand and the Anatolian languages on the other can be seen as a reduction undergone by the original system, and not as an archaic, original feature itself, a feature antecedent to the formation of a richer modal system. The category of *Aktionsart* instead is very well represented across the entire spectrum of the old IE languages. Nevertheless – it must be added – the level of transparency of the *Aktionsart* stems ('maximum transparency': inflectional categories encoding the very notions of 'intensive', 'causative', etc. *vs* 'minimum transparency': inflectional categories denoting mainly aspectual and temporal notions, whilst also subsuming *Aktionsart* notions) varies deeply from language to language. This variation ranges from full grammaticalization (at one end of the scale) toward full lexicalization of the notions in question (at the other end). Finally, the voice is marked through inflectional endings in all those languages that have preserved the opposition: active *vs* middle(-passive). The same holds for the encoding of person and number (with the early disappearance of the dual). The Perfect has not been taken into account here because, as argued elsewhere (Di Giovine 1996: 251 ff.), it represented one of the various, possible *Aktionsart* stems; therefore its presence in a few languages only (Indo-Iranian and Greek, with just a few traces in Germanic and Latin) does not appear to have any major bearing within the genealogical, diachronic framework under discussion.

Let us now go back to the question raised at the beginning of the paragraph, that is: can we draw some general conclusions regarding the unitary nature, or otherwise, of the IE family – as well as its internal articulation – from the analysis of the overall verbal structure (including functional and inflectional categories), as expounded here? Certainly, whatever the answer, this must be cautiously formulated, since genetic correlations are to be established on the basis of a comprehensive examination and consideration of at least the remaining areas of morphology (in particular the nominal inflection), and phonology. Bearing this in

mind, it is evident that the verbal morphology must have constituted a rather compact system across IE, at least at an older stage of development. In fact, the reconstruction of the verbal system of the single IE branches highlights how they get closer and closer to each other as we proceed backward toward the oldest, recorded stages of IE (or to those stages immediately preceding the first records). We can assume therefore the existence of a compact, unitary verbal system, but at a rather shallow, pre-historical time-depth – the fact that we are dealing with a shallow time-depth means that the reconstruction in question is reliable, contrary to what would happen if a reconstruction were itself based on another *reconstructum*[23]. Thus, if we were to judge the status of the IE family from the verbal morphology only, we could not deny the belonging of the various branches to the same common source, a source originally displaying only marginal differences.

At this point there is another, relevant remark to make, a remark that addresses the issue of the evolutionary tendencies, the *drift*, which various IE languages have undergone. This is in fact an area where, in contrast to what stated above, we do observe noticeable divergences: on the one hand there are branches that preserve the original system with hardly any variation (mainly Greek and Indo-Iranian); on the other hand there are (mainly western) languages that present a deep restructuring of the basic configuration of the original system, through a strong tendency to substitute the category of Aspect with that of Tense, as well as through a considerable change in the voice (the passive, that often results from a process of development of the middle, tends now to be encoded through periphrastic constructions with the auxiliary verb, rather than through inflectional endings). Hittite also looses the aspectual opposition in favour of the temporal one, and operates a drastic reduction of the modal categories – in this being in line with the other languages just mentioned – whilst being also characterised by a richness in the *Aktionsart* stems (in this way following the opposite pathway of development undergone by

Greek, for example). Such a profound difference observable not in the original verbal system as such, but in its various *drifts*, cannot be accounted for by modelling the reconstructed IE as a monolithic block. The model that can best account for the evidence is instead that of a linguistic area already differentiated into dialects. Within this linguistic area only some of the potential pathways of development offered by a shared, basic structural system have been actually implemented – in a different way by different languages – whilst others have soon been lost (also through processes of syncretism). This conclusion appears to fit in with another, easily identifiable feature of the verbal system: the multi-functionality of the morphemes[24]. The very fact that many inflectional endings, for example, cannot be connected with a single, unitary function suggests that each IE language has selected a specific set of functions (also on the basis of the combinations allowed within a given verbal form), out of a wider range of potentially available functions.

8. Conclusion

On the basis of the evidence offered by the analysis of the verbal system only, I conclude that there are no elements, no clues that would lead us to deny either the existence of an IE language family or the belonging of Anatolian to this family. It should be emphasized however that there is enough evidence to suggest that reconstructed IE presents itself as a linguistic area already clearly differentiated into dialects, a linguistic area whose main distinctive feature is a very early *drift* in various, different directions, ranging from a rather conservative end point (the Greek-Indo-Iranian core) to a more evolved one[25]. Whatever the case, what is of the essence in the current discussion – I believe – is to stress the need to go beyond the surface, to reach far into the diachronic depth of each language / language group, before establishing hasty comparisons[26]. This is indeed a methodological approach very close to the heart of the first comparativists. This (in my opinion

sound) approach, however, has been a bit neglected by some modern scholars, dazzled by the novelty of data that appeared at first to open the doors to a revolutionary interpretation of the field[27]. Perhaps, all this has been just a delusion.

Endnotes

[1] On the topic of the family tree model as a description of the IE languages see the comprehensive work by Blažek (2007).
[2] Mycenaean has offered only a modest contribution to the understanding of the verbal system.
[3] For a comprehensive summary see Polomé (1994) and Puhvel (1991).
[4] See Schlerath (1987: 44-45) – among others – for a criticism of this method.
[5] For a robust criticism of the method of Glottochronology see Coseriu (1965) and Lazzeroni (1987: 31-32). Recently, Atkinson, Nicholls, Welch and Gray (2003: 206 ff.) have revisited this approach, through cladograms paying much more attention to the data, but still without achieving more persuasive results.
[6] This problem has been noticed by Baldi (1996: 350-51 and 360-61) in his review of Beekes (1995).
[7] As shown elsewhere (see for example Di Giovine (2004: 6) and related bibliography), the reduplication is to be included among the morphemes closely connected to the root, since, formally, this is nothing more than an expansion of the root itself (the prefixes instead have their own form, totally independent from the root).
[8] See Di Giovine (1995) and Di Giovine-Flamini-Pozza (2007).
[9] With regard to reduplication and *Ablaut* in IE see the remarks by Di Giovine (1996: 73-112 and 156-175), in addition to IE grammars.
[10] See again the remarks put forward by Di Giovine (1999).
[11] See Di Giovine, Flamini and Pozza (2007) for the statistical data and the terminology related to this issue.
[12] See Di Giovine (1996: 55-6).
[13] "The Anatolian languages have participated to that pan-IE process of innovation that has given raise to the main inflectional endings (see for example Hittite *esmi* [...]), but does not appear to have a significant modal system. However, if it is true that the innovation in question has been implemented when the various modal forms were all still present and productive, then the IE tradition preserved in the Anatolian languages must have had a rich modal system. In other words, the IE tradition preserved in the Anatolian languages cannot be older than that preserved in the remaining languages".

[14] Modern English belongs to the "non PRO-drop" type.
[15] See the relevant bibliography in Di Giovine (1997a).
[16] Di Giovine (1996: 160-176).
[17] See Di Giovine, Flamini and Pozza (2007: 49-50).
[18] See Bertinetto's and Delfitto's (2000) remarks at this regard.
[19] This interpretation of the function of the suffix *-ške- / -ška-* appears to be confirmed by a recent study by Cambi (2007: 235 ff.).
[20] Wider coverage of this issue is to be found in Di Giovine (2008: 19-21) as well as in Barton (1985: 17 ff) and García-Ramón (2002).
[21] With regard to the issue of the autonomous character, or otherwise, of the dual in IE see Fritz (2000: 133 ff.). In Tocharian the dual is present only in the 3rd person and is restricted to Tocharian B.
[22] See Jasanoff (2003: 26 ff. and 149).
[23] See above § 2.
[24] See Di Giovine (1999).
[25] Obviously, in analysing and explaining the various *drifts* one must take into account also the effects of the contact with other, often non IE languages.
[26] With regard to Anatolian see the excellent methodological remarks pointed out by Carruba (1976: 121 ff.).
[27] See Villar (1979: 171 ff.) for an accurate – though no longer up-to-date – survey of this 'New look' of IE.

References

ATKINSON, Quentin, NICHOLLS, Geoff, WELCH, David & GRAY, Russell
2003 From Words to Dates: Water into Wine, Mathematic or Phylogenetic Inference? *Transactions of the Philological Society* 103/2: 193-219.

BALDI, Philip
1996 Comparative-Historical Indo-European Linguistics: Old and New. *Diachronica* 13/2: 347-364.

BARTON, Charles R.
1985 PIE **su̯ep-* and **ses-*. *Sprache* 31/1: 17-39.

BEEKES, Robert S. P.
1995 *Comparative Indo-European Linguistics: An Introduction.* Amsterdam/Philadelphia: Benjamins.

BERTINETTO, Pier Marco & DELFITTO, Denis
2000 Aspect vs. Actionality: Why they should be kept apart. In Ö. DAHL (ed.). 189-225.

BLAŽEK, Václav
2007 From August Schleicher to Sergei Starostin: On the Development of the Tree-Diagram Models of the Indo-European Languages. *Journal of Indo-European Studies* 35, 1/2: 82-109.

BRUGMANN, Karl
1913 *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, 2nd edn., II/3. Strassburg: Trübner.

CAMBI, Valentina.
2007 *Tempo e Aspetto in ittito*. Alessandria: Edizioni dell'Orso.

CARRUBA, Onofrio
1976 Anatolico e indoeuropeo. In *Scritti in onore di Giuliano Bonfante*, I, Brescia: Paideia. 121-146.

COSERIU, Eugenio
1965 Critique de la glottochronologie appliquée aux langues romanes. In Straka Georges (ed.), *Actes du Xème Congrès International de Linguistique et Philologie Romanes*. Paris: Klincksieck. 87-96.
1997 *Linguistica del testo*. Italian edition by D. Di Cesare. Roma: Carocci.

COWGILL, Warren
1974 More Evidence for Indo-Hittite: the Tense-Aspect Systems. In L. Heilmann (ed.), *Proceedings of the Eleventh International Congress of Linguists* (Bologna-Florence, Aug. 28 - Sept. 2, 1972). Bologna: Il Mulino. 557-570.

DAHL Östen (ed.)
2000 *Tense and Aspect in the Languages of Europe*. Berlin /New York: Mouton de Gruyter.

DI GIOVINE, PAOLO
1995 Sul preterito a vocalismo radicale lungo nelle lingue indoeuropee 'occidentali'. *Linguistica Baltica* 4 = *Kuryłowicz Memorial Volume* (*Analecta Indoeuropaea Cracoviensia*, II). Cracow: Universitas.115-129.
1996 *Studio sul perfetto indoeuropeo. Parte II: La posizione del perfetto all'interno del sistema verbale indoeuropeo*. Roma: Il Calamo.
1997a Aspetto e tempo nel sistema verbale indoeuropeo. In Riccardo AMBROSINI, Maria Patrizia BOLOGNA, Filippo MOTTA &

Chatia ORLANDI (eds) *Scribthair a ainm n-ogaim. Scritti in memoria di Enrico Campanile*. Pisa: Pacini. 309-332.
1997b Gli studi sul sistema verbale indoeuropeo ricostruito: problemi di metodo e recenti prospettive di ricerca. *Incontri linguistici* 20: 11-27.
1999 Funzione e forma nei morfemi e nelle categorie flessionali del sistema verbale i.e. ricostruito. In Jürgen HABISREITINGER, Robert PLATH & Sabine ZIEGLER (eds), *Gering und doch von Herzen. 25 indogermanische Beiträge Bernhard Forssman zum 65. Geburtstag*. Wiesbaden: Reichert. 33-46.
2004 Il perfetto indoeuropeo tra endomorfismo ed esomorfismo. In John H.W. PENNEY (ed.), *Indo-European Perspectives. Studies in Honour of Anna Morpurgo Davies*. Oxford University Press. 3-17.
2008 Il verbo indoeuropeo: tra difettività e suppletivismo. In *Categorie del verbo. Diacronia, teoria, tipologia*. Atti del XXXI Convegno della Società Italiana di Glottologia, Pisa, 26-28 ottobre 2006. Roma: Il Calamo. 11-30.
(to appear) Declino di una categoria flessionale: l'intensivo in greco antico, in: PUTZU, Ignazio, PAULIS, Giulio, NIEDDU, Gianfranco & CUZZOLIN, Pierluigi (eds), *La morfologia del greco tra tipologia e diacronia. Atti del VII Incontro internazionale di linguistica greca* (Cagliari, 13-15 settembre 2007). Milano: Franco Angeli.

DI GIOVINE, Paolo, FLAMINI, Sara & POZZA, Marianna
2007 Internal Structure of Verbal Stems in the Germanic Languages. In Paolo RAMAT & Elisa ROMA (eds), *Europe and the Mediterranean as linguistic areas: Convergences from a historical and typological perspective*. Studies in Language Companion Series, vol. 88. Amsterdam: Benjamins. 49-62.

DINI, Pietro U.
1997 *Le lingue baltiche*. Firenze: La Nuova Italia.

DRESSLER WOLFGANG
1968 *Studien zur verbalen Pluralität - Iterativum, Distributivum, Durativum, Intensivum in der allgemeinen Grammatik, im Lateinischen und Hethitischen*, Wien: Österreichische Akademie der Wissenschaften (= "SbÖAW", 259/1).

FRITZ, Mathias
2000 Der urindogermanische Dual – eine Klasse für sich? In Michaela OFITSCH & Christian ZINKO (eds), *125 Jahre Indogermanistik in Graz*. Graz: Leykam. 133-137.

GARCÍA-RAMÓN, José Luis
2002 Zu Verbalcharakter, morphologischer Aktionsart und Aspekt in der indogermanischen Rekonstruktion. In Heinrich HETTRICH (ed.), *Indogermanische Syntax. Fragen und Perspektiven*. Wiesbaden: Reichert. 105-136.

HOFFMANN, Karl
1967 *Der Injunktiv im Veda*. Heidelberg: Winter.
1970 Das Kategoriensystem des indogermanischen Verbums. *Münchener Studien zur Sprach-wissenschaft* 28: 19-41.

HOFFNER, Harry & MELCHERT, H. Craig
2002 A Practical Approach to Verbal Aspect in Hittite. In Stefano DE MARTINO & Franca PECCHIOLI DADDI (eds), *Anatolia Antica. Studi in memoria di Fiorella Imparati*. Firenze: Lo Gisma. 377-390.

JASANOFF, Jay H.
2003 *Hittite and the Indo-European Verb*. Oxford University Press.

LAZZERONI, Romano
1977 Fra glottogonia e storia: ingiuntivo, aumento e lingua poetica indoeuropea. *Studi e Saggi linguistici* 17: 1-30.
1987 Il mutamento linguistico. In R. LAZZERONI (ed.), *Linguistica storica*, Roma: La Nuova Italia Scientifica.
1995 Sulle orme di Vittore Pisani: l'indo-ittita e la questione dei modi indoeuropei. *Paideia* 50: 225-231.

MEID, Wolfgang
1975 Probleme der räumlichen und zeitlichen Gliederung des Indogermanischen. In H. RIX (ed.), *Flexion und Wortbildung. Akten d. V. Fachtagung d. Indogermanischen Gesellschaft*. Wiesbaden: Reichert. 204-219.
1979 Der Archaismus des Hethitischen. In E. NEU & W. MEID (eds), *Hethitisch und Indogermanisch*. Innsbruck: Institut für Sprachwissenschaft der Universität (= IBS, 25). 159-176.

MEILLET, ANTOINE
1937 *Introduction à l'étude comparative des langues indo-européennes*, 8ème ed., Paris, repr. Forge Village Mass.: University of Alabama Press.

MELCHERT, H. Craig
1997 Traces of a PIE Aspectual Contrast in Anatolian? *Incontri linguistici* 20: 83-92.

NEU, Erich
1985 Das frühindogermanische Diathesensystem. Funktion und Geschichte. In B. SCHLERATH (ed), *Grammatische Kategorien - Funktion und Geschichte. Akten d. VII. Fachtagung d. Idg.*

Gesellschaft (Berlin, 20.-25. Feb. 1983). Wiesbaden: Reichert. 275-295.
1989 Dichotomie im grundsprachlichen Verbalsystem des Indogermanischen. *Indogermanica Europaea - Festschrift für Wolfgang Meid zum 60. Geburtstag am 12.11.1989*. Graz: Institut für Sprachwissenschaft. 153-175.

PINAULT, Georges-Jean
1989 Introduction au Tokharien. In *Lalies - Actes des sessions de linguistique et de littérature*, VII (27 août - 1er septembre 1985). Paris: Presses de l'École Normale Supérieure. 5-224.

POLOME, Edgar C.
1985 How Archaic is Old Indic? In U. PIEPER & G. STICKEL (eds), *Studia Linguistica Diachronica et Synchronica W. Winter Sexagenario ... Oblata*. Berlin: Mouton de Gruyter. 671-683.
1994 Indo-European Studies since the Neogrammarians. In G. E. DUNKEL, G. MEYER, S. SCARLATA & Chr. SEIDL (eds), *Früh-, Mittel-, Spätindogermanisch, Akten der IX. Fachtagung der Indogermanischen Gesellschaft vom 5. bis 9. Oktober 1992 in Zürich*. Wiesbaden: Reichert. 295-314.

PUHVEL, Jaan
1991 Whence the Hittite, Whither the Jonesian Vision? In S. M. LAMB & E. D. MITCHELL (eds), *Sprung from Some Common Source - Investigations into the Prehistory*. Stanford Ca.: Stanford University Press. 51-66.

SCHLERATH, Bernfried
1987 On the Reality and Status of a Reconstructed Language, *Journal of Indo-European Studies* 15, 1/2: 41-46.

SCHMITT-BRANDT, Robert
1987 Aspektkategorien im PIE? *Journal of Indo-European Studies* 15, 1/2: 81-92.

STRUNK, Klaus
1994 Relative Chronology and Indo-European Verb-System: The Case of Present- and Aorist-Stems. *Journal of Indo-European Studies* 22, 3/4: 417-434.

STURTEVANT, Edgar H.
1929 The Relationship of Hittite to Indo-European, *Transactions and Proceedings of the American Philological Association* 60: 25-37.

SZEMERÉNYI, Oswald
1990 *Einführung in die vergleichende Sprach-wissenschaft*, 4. Aufl. Darmstadt: Wissenschaftliche Buchgesellschaft.

VILLAR, Francisco
1979 Hetita e indoeuropeo, *Emerita* 47/1: 171-188.

Stratified Reconstruction and a New View of the Family Tree Model

Bridget Drinka
University of Texas, San Antonio

In this article, three related arguments are presented: first, that the family tree model as it stands is an inadequate depiction of the complex relationships of the Indo-European languages, and that a new model is needed; secondly, that the stratification of archaic vs. innovative structures allows us not only to recognize the Indo-European languages as related, but also to acknowledge that some Indo-European languages, like Indo-Iranian and Greek, must have remained in contact longer than others; and, finally, that this late contact constitutes one more piece of evidence that the 'Out of India' theory, whether referring to the indigeneity of Indo-Aryan on the Indian subcontinent or the origination of proto-Indo-European itself in that location, is untenable.

0. Introduction

Recent debates[1] concerning the proposed indigeneity of the Indo-Aryan languages to the Indian subcontinent have generated a number of related controversies: should the Proto-Indo-European *Urheimat* be located in India, as the "Out of India" theory would require? Can linguistic data trump archeological data in the reconstruction of prehistoric cultural relations and developments? Are the comparative method and the genetic model outdated and unusable, having been superseded by other explanations, such as language contact?

This paper starts from the assumption that linguistic evidence is, indeed, an indispensable resource in cultural reconstruction, one which provides more complete information as to the relatedness of particular groups than the archeological record

can furnish. The validity of Proto-Indo-European as a construct is not contested here, nor is the importance of comparative reconstruction as a valuable tool. What is contested, however, is the over-simplistic characterization of linguistic relationship and change still maintained by a number of Indo-European scholars. Here I explore three related controversies and the implications which each of them entails:

1. The claim that linguistic relationship can be depicted by a simple family tree is strongly contested. The genetic model is here endorsed as a descriptive model of the ancient relationship of Indo-European languages, but its usefulness as an *explanatory* model is questioned. A new model, one which incorporates both genetic and geographical information, is proposed. The importance of areal spread, of the "horizontal" nature of innovation, is explored in some detail.

2. This horizontal nature of change implies a stratification of data, a layered distribution of archaic and innovative features. Hittite and Germanic, for example, appear to have separated from the IE unity at an earlier time, judging from their archaic verb systems and other evidence; Greek and Indo-Iranian, on the other hand, share a number of morphological and other innovations, and seem to have remained in contact longer.

3. This stratification, in turn, tells us something about the probability of the "Out of India" theory: besides other arguments that can be adduced, the late contact of Indo-Iranian and Greek reinforces the view that a development of Indic or of Indo-European on the Indian subcontinent is completely untenable.

1. Models of Relationship

1. 1 Trees and the representation of genetic relationship

The use of trees to represent the relationships among IE languages dates to the early days of Indo-European studies: Schleicher's renowned *Stammbaum* (1860) provided a foundational image of genetic relatedness among IE languages.

While the limitations of the family tree model have been recognized virtually from the time of its inception (cf. Schmidt's *Wellentheorie* (1872), proposed as a counterpart to the family tree model), it has remained as a simple organizational principle through the years.[2] Recently, new work on cladistic models in biology and improvements in computer technology have led to a renewed interest in the family tree as a viable representation of descendency and relatedness among languages. The research team of Don Ringe, Ann Taylor, Tandy Warlow, and collaborators (e.g., Taylor et al. 2000; Ringe et al. 2002; Nakhleh et al. 2005) has attempted to establish a "perfect" phylogenetic reconstruction of Indo-European and other language families. These studies, taken together, provide an informative illustration of how cladistic models, similar to those used in genetics, are being constructed to represent linguistic relationships. A careful look at the first two of these studies will allow us to examine the validity of the model and implications of the arguments presented there.[3]

The preliminary study of Taylor et al. (2000) lays out the premises and procedures of the tree-construction enterprise: for a given group, particular features ("characters")—phonological, morphological, or lexical[4]—are coded and fed into an algorithm which produces a tree. Borrowing and parallel developments are to be excluded, since these "do not reflect the genetic descent of the languages involved" (Taylor et al. 2000: 397). The perceived need to exclude borrowed forms leads the researchers to focus on phonological and morphological characters, since lexical items are much easier to borrow. Two sound changes, the "satem" development[5] and the "ruki" rule[6], are given special prominence. The trees which result depict the "satem" languages of Vedic, Avestan, Old Church Slavonic, and Lithuanian as forming a genetic subgroup which they call the "satem core" (2000: 400-3). In the west, Germanic was not so easily positioned on the tree, in some ways aligning itself with Balto-Slavic, in other ways with Celtic and Italic. The researchers suggest that Germanic began as the "nearest sister" of Balto-Slavic, but switched affiliation by

means of contact with Celtic and/or Italic. Germanic, then, is depicted on the "non-tree edges," as in Figure 1.

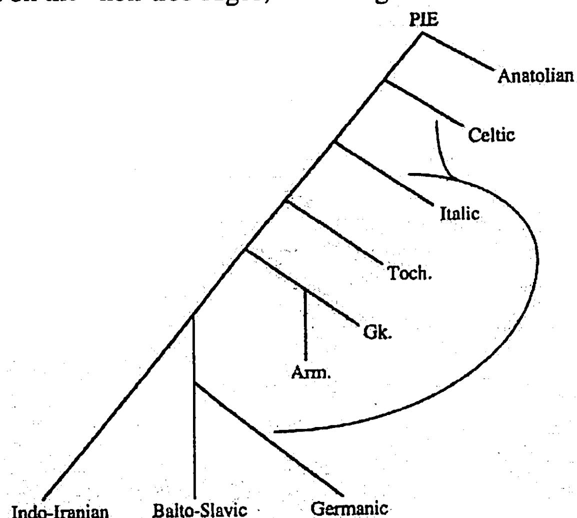


Fig. 1: The complex position of Germanic (from Taylor et al. 2000: 407) (with kind permission by John Benjamins)

Several problems with this analysis are immediately evident. First of all, while lexical items are, indeed, more easily borrowed, phonological and morphological properties can certainly also spread from language to language, as has been amply illustrated in the literature on languages in contact (Thomason and Kaufman 1988; Thomason 2001; Drinka 2003a; Heine and Kuteva 2005, 2006). And, ironically, the two sound changes that the researchers use for their analysis, the "satem" development and the "ruki" rule, have frequently been viewed as examples of language contact, with the effects of these innovations diffusing across a broad area of Indo-European languages. But the trees presented in Taylor at al. (2000: 400-3) depict the "satem" languages under a single node, as if they formed a genetic subgroup. This grouping together of Indo-Iranian and Balto-Slavic gives the false

impression that these languages shared a common ancestor, rather than acquiring features which spread across the territory.[7]

An additional problem is that, in some iterations of the algorithm (especially when Germanic is excluded), Old Irish and Latin are also linked under such a node, recognizing the phonological and morphological similarities that these languages share, but depicting this similarity, again, as a sign of genetic relationship rather than more credible late contact.[8]

Finally, as depicted in Figure 1 above, they characterize the relationship that Germanic had with other Indo-European languages as sharing a node first with Balto-Slavic (as its "nearest sister") and then with Italic and Celtic. Because of the need to force the data into tree structures, they fail to recognize that these connections were clearly areal in nature, as demonstrated in some detail by Porzig (1954).[9]

Ringe et al. (2002) expand upon Taylor et al. (2000), but continue in the same vein, giving primacy to genetic explanations over those involving contact. The authors argue that normal transmission of language does not lead to mixed grammars, but that such mixings can only result from discontinuities of transmission—an occurrence which they judge to be rare (Ringe et al. 2002: 63). In their estimation, contact is to be virtually excluded from historical analysis.[10]

One of the important criteria they use to construct the IE family tree is the organization of the verb system in the various IE languages (M1) (Ringe et al. 2002: 117-19). The authors assign languages to the following categories according to the distribution of their verbal structures:

1. one stem per lexeme
    a. two conjugations [Hitt.]
    b. single conjugation [Luv., Lyc.]
2. present / aorist / perfect contrast
    [Arm., Gk., Alb., Ved., Av., OCS, Lat., OPers., Osc., Umb.]

3. present / subjunctive / preterite contrast, the former two largely parallel [TA, TB]
4. present / preterite / infinitive contrast [Lith., OPrus.]
5. present / preterite contrast, the latter in two conjugations ("strong" vs. "weak") [OE, Goth., ON, OHG]
6. present / subjunctive / future / preterite contrast [OIr.]
7. present / subjunctive / preterite contrast, the latter two usually sigmatic [Welsh]

There are several questions that need to be asked with regard to this method of judging relationship: we need to know if all of these distributions are viewed as distinct, single instantiations of change, or whether a number of separate innovations are being posited as responsible for the final configuration. For example, should the presence or absence of a subjunctive construction separate languages which are otherwise parallel? Subjunctives are closely aligned with other parts of the verb system, frequently developing from present or future tense verbs in subordinate clauses (Bybee et al. 1994: 230-36). It seems preferable to consider the development of the moods separately, as the authors do with the thematic optative (M6).

On a deeper level, the above characterization of verbal organization points to a problem with the methodology. The arrangement of the categories sets up the apportionment of the languages into particular predetermined categories, but fails to identify the relationship of one category to another, or to establish how these distributions reflect the nature of the verb system in the proto-language. It appears that the prime goal in this exercise is to establish relationship among daughters, and not to determine which systems retain archaic distinctions more faithfully, and which have innovated. For example, why would the sigmatic nature of the subjunctives and preterites of Welsh be singled out for consideration, and not those of, say, Greek? Is the claim being made that sigmatic forms are only crucial in Welsh?

Along similar lines, there seems to be no recognition of the fact that some verbal categories are clearly built upon others. For example, the root aorists represent a very archaic layer in the IE verb system; many present stems are built upon these old root forms by the addition of infixes and suffixes, and the imperfects are, in turn, built upon these present stems by the affixation of the augment (past prefix) and secondary (past) endings:

Skt. √*kṛ* 'make':

Root Aorist: *akaram*	Imperfect: *akṛṇavam*
a- *kar-* *am*	*a-* *kṛṇav-* *am*
aug. + aor.stem + sec.end.	aug. + pres.stem + sec. ending (*kr* + *ṇ*-infix)

Ringe et al. do set up a nesting of branches, implying that particular stages of a language—as represented by nodes—precede others in time. While archaism and innovation in the verb system are not altogether excluded from consideration (cf. the commendable inclusion of the thematic aorist (M3) as an indicator of innovation [Ringe et al. 2002: 94]), the telling clues from stratified temporal and aspectual categories are largely ignored.

As a result of inattention to such details, this system does not allow us to recognize the remarkable similarity between a number of features of the Greek and Indo-Iranian verb morphologies, such as their identical formation of the long-vowel subjunctives and their very similar expansion of reduplication in the perfect systems. The authors do recognize that these two branches alone, along with Armenian, have adopted the augment (M2), but they do not mention the far-reaching implications of this adoption: the addition of the augment enabled both Greek and Indo-Iranian to form an imperfect alongside their older aorists, as illustrated above. In this way, these languages developed a productive aspectual system, which contrasts to the older, simpler system such as that found in Hittite and Germanic (present vs. preterite). What is left unnoticed is that the innovative use of the

augment also appears to be areally distributed, not unlike the "satem" rule and the "ruki" rule, though encompassing different languages and occurring at a different time. The shared expansion of the verb system in Indo-Iranian and Greek, then, is not focused upon because the characters which mark it are being used to determine filiation rather than shared innovation. In the last analysis, this late, areal relationship between Greek and Indo-Iranian should not be placed on a tree.

In summary, there are at least four related points that can be made concerning the suitability of the family tree model to represent the facts of Indo-European and other language families:

1.2.1. Family trees are non-explanatory

While there is clear value in charting proposed relationships among languages, we must also admit that the family tree model represents one perspective, among several possibilities, organizing the data so that we may detect archaism, and projecting these archaisms back to a predecessor language. The point of constructing a tree is not to sort out why languages developed as they did. It is to ignore these motivations, to weed out evidence of what occurred in the ensuing years, in order to arrive at a point of unity in the past. What the family tree represents, then, is a skeleton, providing a mapping of common ancestry, but no information as to how or why the languages diverged. A family tree does not help us understand why particular features changed and why others remained intact;[11] stasis is thus regarded as an un-analyzable fact, just as innovation is. It is, of course, the case that the family tree was not designed to account for change. Trees are, by their nature, purely descriptive. However, if we seek a true characterization of the Indo-European family as it has become, we need a more comprehensive model, one which addresses how and why innovations entered the daughter languages, one which takes into consideration the extent to which these innovations were adopted or not adopted by other varieties due to reasons other than shared inheritance. It is important to note that, when speakers

adopt change, they are not doing so because of their language stock, that is, because of those to whom they are related, linguistically. They are adopting change because of those with whom they are interacting, and with whom they wish to align themselves.

1.2.2. Family trees are inaccurate depictions

As Ringe et al. themselves note (2002: 109), trees may not accurately represent the existence of dialect continua: when only certain varieties of a dialect continuum survive, analysts may mistakenly judge them to represent "clean speciation."[12] A tree may then represent "not the original diversification of the languages but the end product of a complex series of events of differentiation and survival." Only the survivors are recognized—a remnant of political or social success more than of purely linguistic development.

1.2.3. Family trees undervalue contact as an essential element of change

By forcing data onto trees, Ringe et al. disallow the possibility that a change has spread across the landscape. For example, although they do admit (Ringe et al. 2002: 109), as mentioned above, that the "satem" innovation could have spread from Indo-Iranian into Balto-Slavic after a time of unity, they still use this innovation to establish a closer *genetic* relationship between Indo-Iranian and Balto-Slavic than, say, between either of these and Greek (See Fig. 2).

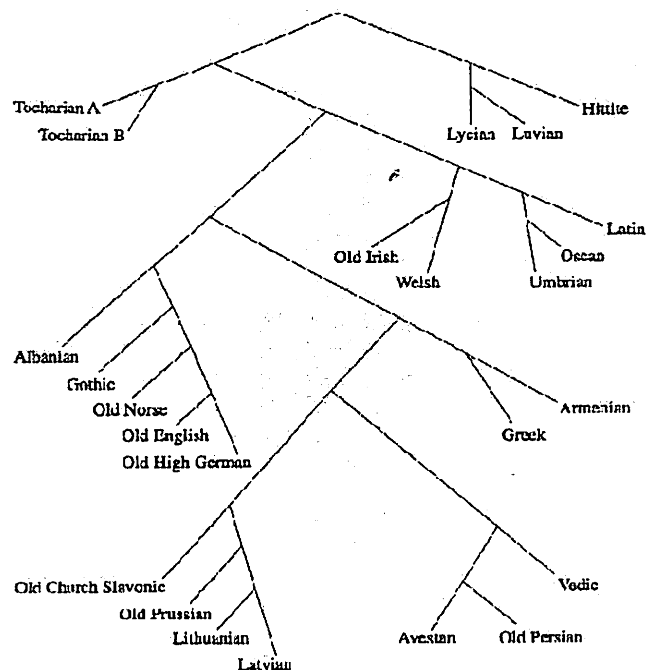


Fig. 2: The apparent best tree for the entire Indo-European dataset (Ringe et al 2002: 87). (Reprinted by permission of Blackwell Publishing)

Their model simply does not have a mechanism to account for diffusion. Of course, it must be admitted that, just as the family tree model was not designed to explain change, it was also not meant to account for areal spread or contact phenomena either, as McMahon & McMahon (2005: 18) note:

> We must acknowledge that family trees cannot tell the whole story, but equally that they do capture one important aspect of linguistic history; this does not mean we should castigate or reject the tree model for not incorporating contact, which it was never designed to do in the first place.

Nevertheless, when morpho-syntactic change due to contact is excluded because it does not fit the model, we are left to wonder if

the inadequacy of a model should lead to the elimination of recalcitrant data from consideration. A preferable approach, it would seem, would be the creation of a more comprehensive model which could represent other possible scenarios of development in addition to the genetic one.[13]

1.2.4. Family tree models are not sufficiently complex.

Recent work on cladistics in Indo-European for the most part replicates and reinforces work that has already been done, notably by scholars like Porzig (1954), Birwé (1955), and Euler (1979), who sorted out many of the relationships of IE languages, both those due to retention of archaism and those representing shared innovation, in fairly minute detail. For example, Nakhleh et al. (2005: 189) regard the distribution of the medio-passive **-r*, found in Anatolian, Tocharian, Italic, and Celtic, vs. **-y /-i*, found in Germanic, Greek, and Indo-Iranian, (M5), as representing a split of these languages into two branches. Porzig (1954: 85), on the other hand, following Meillet (1931: 5), notes the areal distribution of the innovative **-i* form, which suggests a "peripheral" vs. "central" distribution of the two categories as more correct, and not as indicative of genetic grouping; (see also Drinka 1993: 418-25; 1999: 490).

I would like to suggest that we can do more with our computational skills than reconstruct skeletal trees of archaic retentions. Rather than excluding contact evidence as ancillary, we should be finding ways to incorporate evidence from linguistic contact into our depiction of change within the IE languages. McMahon and McMahon (2005) present an array of potential computational analyses, and regard the genetic tree model as useful but in some ways inadequate. They point to Kessler's (2001) suggestion that, while a strict adherence to cognate sets could still be used for tree construction, analysts should also make use of a fuller data set, including linguistic loans, when "historical connectedness" is sought, i.e., for comparison with archeology or

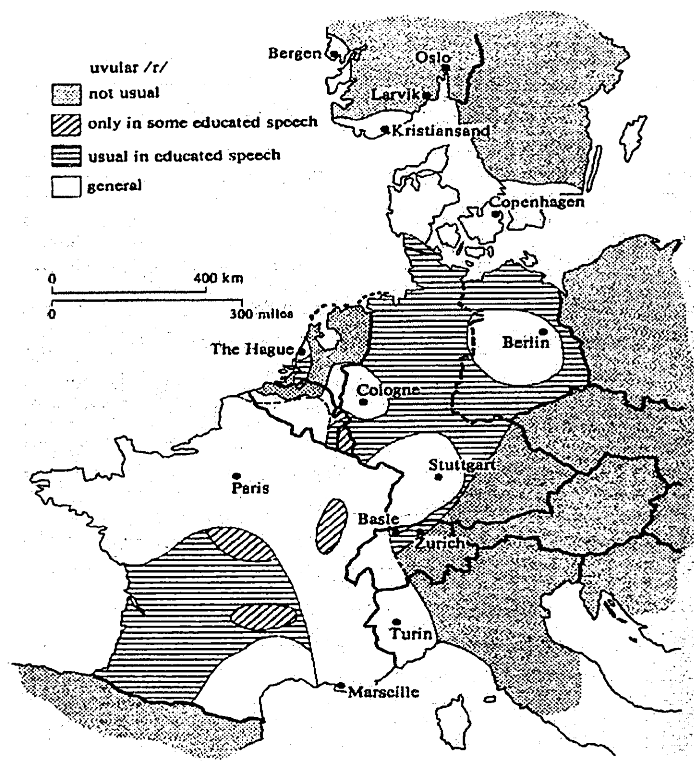
genetics. McMahon & McMahon (2005: 137) go on to endorse the incorporation of contact data:

If we are serious about rehabilitating contact-induced change and want to be able to account for both aspects of Kessler's 'historical connectedness' (2001), then our concentration on trees is problematic.

1.3 Waves, rivers, and other acladistic models of linguistic change

In response to and as a supplement to Schleicher's *Stammbaumtheorie*, Johannes Schmidt proposed the *Wellentheorie*, or Wave Theory (Schmidt 1872): change was viewed as emanating from a central point, like waves spreading from the spot where a stone is dropped into a pool. Areas located nearest to the central point are expected to adopt the change most thoroughly; more distant locations may undergo the change only partially, in particular environments.

While appealing in theory, the model has been shown to be somewhat inaccurate in predicting the actual movement of innovations across a landscape: as Trudgill (1983: 58-9) demonstrates, innovations, at least those which occur in modern times, do not move in steamroller-fashion across the countryside, but jump from one large urban center to the next largest (though not necessarily contiguous) urban center, skipping over the intervening territory. For example, as illustrated in Map 1, the uvular *r* of Paris spread first to Berlin, Torino, and Brussels, not to smaller Osnabrück, Bergamo, or Luxembourg.



Map 1: The distribution of uvular *r* (Trudgill 1983:58). (Reprinted by permission of Blackwell Publishing)

This "gravity model," illustrated in Fig. 3, thus factors in population size, along with the distance from the point of inception.

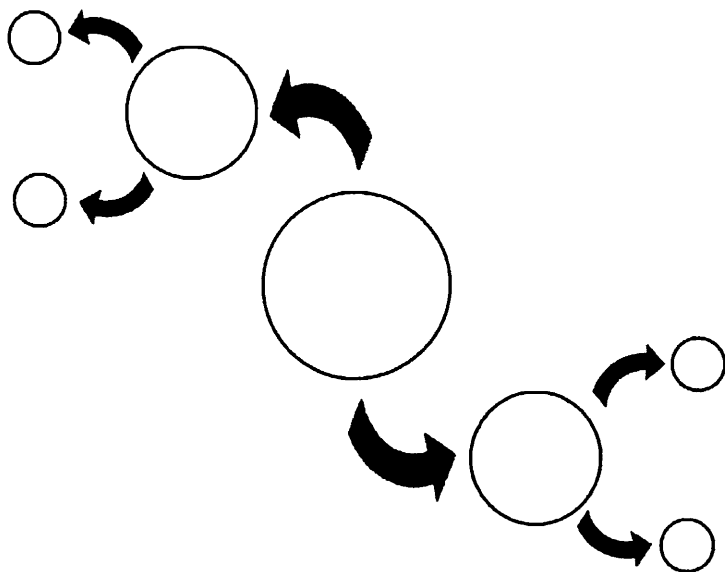


Fig.3 The gravity model of linguistic diffusion (based on Wolfram & Schilling-Estes 2003:724).

Mufwene (2001) also provides several alternative models to the family tree to characterize the spread of innovations across a population. He makes an intriguing comparison of language to a parasitic species, which depends upon its host (the speakers) and society to survive (Mufwene 2001: 150):

> ...linguistic features are passed on primarily horizontally, more or less on the pattern of features of parasites, through speakers' interactions with member of the same communicative network or of the same speech community.

Initial agents of change introduce linguistic features from one community to another, just as they would germs (2001: 151). Mufwene also proposes a less fully-articulated yet extremely apt model of a river as an analog for linguistic change (Mufwene

2001: 143): as a stream flows along, it is colored by the various tributaries which flow into it; it is affected by contact with these other bodies of water. As the river reaches its delta and splits into several smaller streams, each retains remnants of the larger stream, but each also assumes the coloration of the sands beneath. In similar fashion, we can envision language as it moves through time and space, affected by its own ecological context—in this case, contact with other varieties, sub-stratal influences, etc.—all of which influence the outcome.

Granted that innovations move across populations in horizontal fashion, should the family tree model be abandoned in some locations (e.g., Australia, as argued by Dixon 1997, or in the Uralic area, as argued by Marcantonio 2002), or given up altogether, in favor of a wave-like model? Most scholars recognize some value in retaining information about genetic relatedness when that information is reconstructable, but a number of researchers protest that the tree as it is currently conceived is not the appropriate depiction of this relatedness. Harrison (2003), for example, claims that, while relatedness can be established by the comparative method, the actual image of this relatedness is not well represented by the sub-groupings that appear in family trees. In his view, innovations "diffuse through the linguistic landscape, and give rise to the patchwork of isoglosses rather than the discreteness of trees" (2003: 239).

With regard to the appropriateness of trees for the analysis of Indo-European, Garrett (1999:152-3) takes a rather similar acladistic tack to that of Harrison in his model of IE sub-grouping and dispersal. He posits a non-distinct predecessor to Celtic, Italic, and Greek and proposes that the languages which eventually developed may have done so through "local responses to areal and cultural connections," that is, in the context of "secondary Sprachbund phenomena." He goes on to develop these claims more explicitly in Garrett (2005), stating that the tree-like image of IE phylogeny should be given up in favor of an image of "convergence among neighbouring dialects in a continuum"

(Garrett 2005: 139). Dialects, then, would not have been well-formed "Celtic" or "Italic" languages when they shared geographical space in early times. Garrett depicts the relationship as a triangle (Fig. 4), still depending upon the structure of a tree, but without the usual lines and nodes of direct descendency.

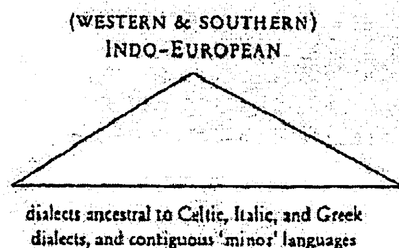


Fig. 4: Model of western and southern IE sub-grouping (Garrett 1999: 152).

1.4 Amalgamation of "Tree" and "Wave"

Calvert Watkins (2001: 63), in his exploration of areal influence in ancient Anatolia, claims that, while the comparative method is usually confined to the analysis of genetic relation, it can also be used to analyze areal diffusion, as well:

> The goal of genetic comparison is linguistic history, while that of typological comparison is often said to be linguistic universals. But one can and, I insist, must compare the components and manifestations of a linguistic area in order to draw *historical* conclusions. (emphasis his)

Although he does not give details as to how this method would be implemented for use with linguistic areas, I would argue that Watkins is correct in his assessment: innovations caused through contact must be accounted for in any explanation of historical development. When the goal of the analysis is reconstruction of a proto-language, then the intervening influences and changes are collected in order to be eliminated from consideration. But when the goal of the inquiry is to construct an accurate depiction of the history of the language, to pay attention to those non-ancient, innovative elements that have entered a language, then carefully conducted comparative analysis will help linguists sort out what is shared because of genetic relationship and what is shared because of contact with other varieties.

Chappell (2001: 354) suggests that a more elaborate, synthetic approach is required:

> To reconstruct the history of a language family adequately, a model is needed which is significantly more sophisticated than the family tree based on the use of the comparative method. It needs to incorporate the diffusion and layering process as well as other language-contact phenomena such as convergence, metatypy and hybridization. The desideratum is a synthesis of all the processes that affect language formation and development.

Although Campbell (2006: 19-20) takes issue with this statement, which seems to him to be a challenge to the comparative method, he states something not dissimilar himself:

> Mainstream historical linguists realize that it is not possible to understand diffusion fully without knowing the genetic affiliation of the languages involved, and vice versa, it is not possible to account fully for what is inherited without proper attention to what is diffused. That is, it is not two distinct, opposed and antagonistic points of view that are involved, but rather both are needed and they work in concert. (Campbell 2006: 18)

While a number of scholars, then, have recognized the desirability of analyzing both genetic and contact factors at once, a successful amalgamation of the family tree and wave models, as sought by Schmidt, has not yet been produced. A large step in this direction has been taken, however, by McMahon & McMahon (2005; 2006), who have conducted an extensive analysis of phylogenetic classification systems with an eye to discovering which could best recognize the influence of contact as well as genetic filiation. According to their findings, network models are superior to family tree models, since the former are capable of representing not only shared origin, but also divergence caused by

contact (2005: 178; 2006: 70). Among the most successful network models, they argue, is one used in biology and genetics, NeighbourNet, which constructs either networks or trees, as needed, to represent the similarities and distances between languages (McMahon & McMahon 2005: 158). What seems particularly attractive about this model is that it is, indeed, able to detect the effects of contact, marking these effects by means of "reticulations":

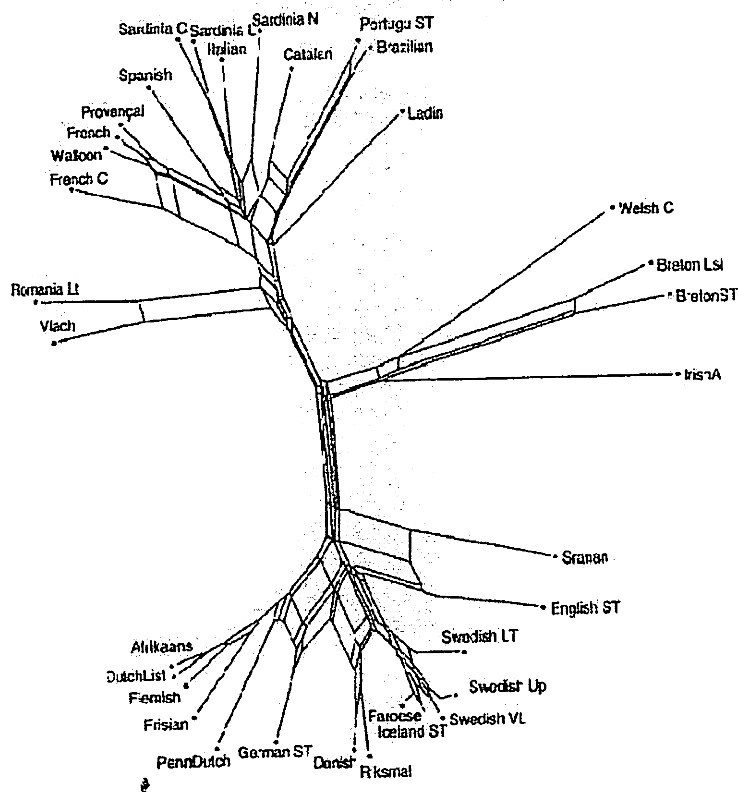


Fig. 5: Germanic and Romance lolo data on NeighbourNet (McMahon & McMahon 2005: 161). (Reprinted by permission of Oxford University Press).

Fig. 5, for example, represents a network produced by NeighbourNet using the least conservative Germanic, Romance, and Celtic lexical data from the Dyen, Kruskal, and Black (1992) database.[14] The reticulations, or box shapes, represent ambiguities in the data that can be interpreted as signs of contact. For instance, Vlach and Romanian are well removed from the other Romance languages, presumably because of the influence of Greek (not shown on this network); the three varieties of Swedish are tightly connected through dialect contact as represented by the density of their reticulations; Sranan, though an English creole, has a clear link here with Dutch, as well, according to the authors, documenting the fact that much lexical borrowing has occurred. Figure 6, which uses the full data set from Dyen, Kruskal, and Black (1992), generates a tree, but also produces a number of reticulations, connoting contact, especially in Balto-Slavic. While reticulations still do not provide information as to how contact operates, they at least signal places where contact has apparently occurred, while also signaling genetic relationship.

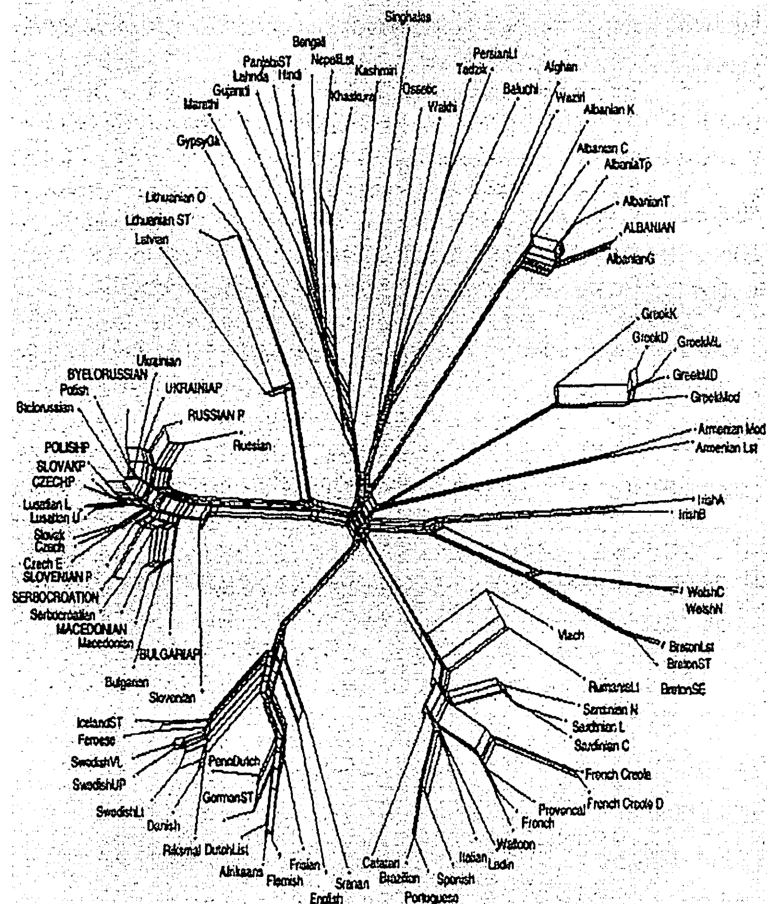
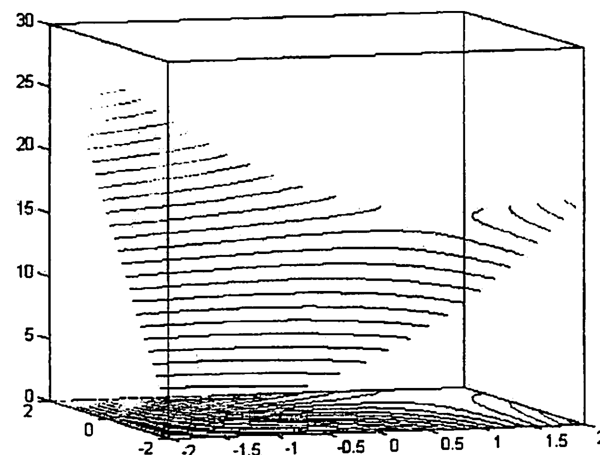


Fig. 6: Indo-European relations using NeighbourNet (McMahon & McMahon 2005: 164). (Reprinted by permission of Oxford University Press).

What I would like to propose is that it should be possible to represent both genetic relationship and diffusion in one model, if we construct a 3-dimensional model. Such a model can be envisioned, in idealized terms, in the following way: Imagine that

every citizen of Eurasia dons a colored cap corresponding to his or her native language—Tamil speakers might wear red, Slovenian green, Estonian white, etc. All would stand outside while a satellite passes over the continent, photographing the entire array. One hundred years later, the same event would take place, and each century thereafter for several millennia. These photographs could then be stacked one atop the other. What would be the result? Let us say that the population of Slovenian speakers expanded over the course of the centuries, spreading into neighboring territories. The stacked 2-dimensional snapshots would document this expansion as a cone of green, starting from a fairly small locality but spreading out as time went on (see Fig. 7). If the language developed sub-varieties over time, these could be represented by bi-or tri-furcations, as illustrated here. Now if we picture turning this cone upside down, we can see something resembling a 3-dimensional branch of a family tree: the small end, basically present-day Slovenia, is like a node, and the (bifurcated) cone which extends from it, is not unlike Garrett's triangle of undefined filiation, but in three dimensions (see Fig. 8).



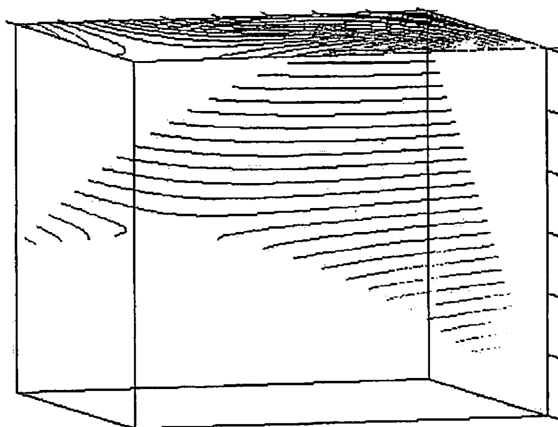


Fig. 7 and Fig. 8: The 3-dimensional, amalgamated model.

What is, I think, appealing about this model is that it allows us to recognize that change—whether it is expansion of languages into new territory or the spread of features across populations—occurs in horizontal fashion, as Mufwene (2001) points out, but that it also allows us to keep track of genetic information through the addition of the vertical element of time.

This sketch is admittedly rough, and in need of refinement: it would be more realistic and practical, for example, to examine the distribution of individual features rather than entire languages, since features do tend to spread to neighbouring varieties through contact, whereas entire languages would be expected to spread primarily when speakers themselves move or when populations of speakers mix. Although the relation to the family tree would be diminished by the focus on features, our understanding of the stratified nature of innovation would be enhanced: each layer of change would be stacked upon previous layers of innovation. The accumulation of these innovations through time gives us our present-day result.

Several other possible refinements of this 3-d model can also be mentioned. First of all, the bilingual or multilingual nature of many speech communities needs to be represented. If each speaker were assigned, say, a pixel on the map for each language he or she used, areal trends would still emerge, and genetic relationship would still be recognizable. Furthermore, some recognition of the "gravity model" needs to be incorporated into this representation (see 1.3 above), recognizing that innovations tend to spread to non-contiguous urban areas, and not in simple wave-like fashion to all surrounding areas. Finally, the argument can be made that this model is still not explanatory: an attempt is being made to illustrate how change occurs, through the spread of innovations, but simply mapping these innovations, even through time, does not tell us why they occur. We must still look to sociolinguistic and other explanations to be able to conclude what actually caused a given change. What I would claim, however, is that such a model at least allows these motivations to be considered.

2. Stratified Models of Indo-European Relationship

Besides providing us with an amalgamation of the "family tree" and "wave" models, then, this approach allows us to envision a "stratified reconstruction" of Indo-European, in which various daughter languages separated from a dynamic "proto-conglomerate" at different points in time. Evidence for the validity of this layered approach, as opposed to the image of the *Ursprache* as a single, unified entity, can be found in such naturally stratified linguistic systems as the temporal, aspectual, and modal categories of the IE verb. Some languages display more archaic morphological features, and so may have separated from the proto-language at an earlier time; others show accretion, innovations built upon the archaic layers, which suggest that these varieties separated later. For example, both Hittite and Germanic have extremely simple verb morphologies, focusing on the temporal

contrast of present and past, with no sign of an ancient subjunctive or future, and no systematic aspectual contrast in the past tense (Polomé 1964; 1982; 1987). Greek and Indo-Iranian, on the other hand, must have remained in contact for a longer period of time. These "eastern" languages possess, among other features, an extremely similar temporal-aspectual system, including a productive 3-way contrast of present-imperfect-aorist, as opposed to the 2-way contrast seen elsewhere in Indo-European; a lengthened theme-vowel subjunctive, not found elsewhere; and an obligatory use of reduplication to mark the perfect, found only sporadically in the west (Birwé 1955; Drinka 1995b & 2003b).[15] The shared morphological complexity of Indo-Iranian and Greek thus appears not to represent archaism, as had traditionally been thought, but rather shared innovation due to late contact in the eastern area. Following Meid (1975), we can represent this stratification of languages across space and time as follows:

Fig. 9: **Model of a developing proto-language** (Meid 1975)

I. *early Indo-European* (c. 5th millennium B.C.) represented by archaisms in both the eastern and western areas
II. *middle Indo-European* (c. 5th-4th millennium B.C.) represented by more recent features found in both east and west
III. *late Indo-European* (3rd-2nd millennium B.C.) represented by recent innovations in differentiated languages
    a) *eastern group:* esp. Greek, Indo-Iranian
    b) *western group*: esp. Italic, Celtic, Germanic

This model sets up a framework of early, middle, and late stages in the proto-language, and posits a separation of eastern and western areas in the last stage, representing the shared changes which occurred separately in the east and in the west.

In a model which integrates archeological and linguistic data, including evidence of ancient linguistic and cultural contacts, David Anthony (2007: 55-56) has taken a step further in

recognizing the important link between Greek and Indo-Iranian noted in Meid's model.[16] His diagram of the sequence of splits by the IE languages (Fig. 10) also portrays a very early separation for Anatolian, at a time before wheeled vehicles would have existed in the steppes, as well as early separations for Tocharian and Germanic, but a later split for Greek, and even later for Indo-Iranian.

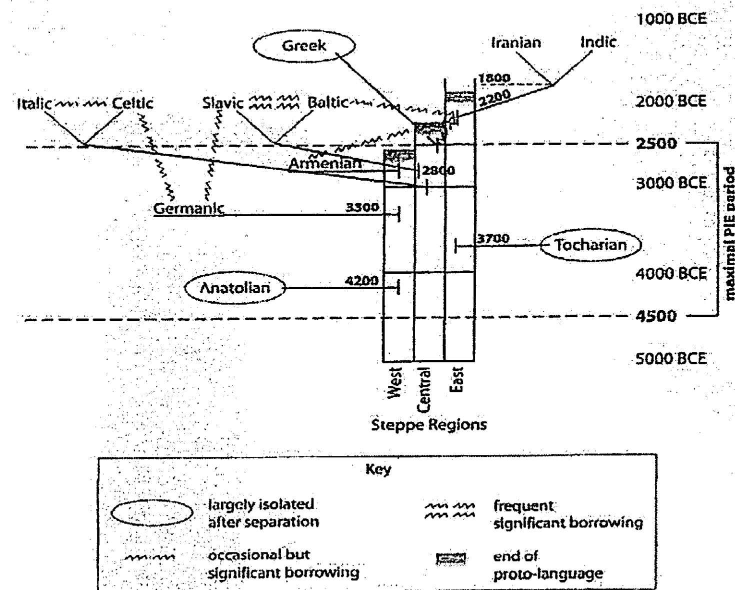


Fig. 10: A diagram of the sequence and approximate dates of splits in early IE (Anthony 2007: 100). (Reprinted by permission of Princeton University Press)

In this model, the contact between Greek and Indo-Iranian implies that Greek did not move out of the central region until fairly late. The fact that Greek did not take part in either the "satem" or the "ruki" innovations, however, provides a *terminus ante quem*: Greek and Indo-Iranian were in contact late, but not as late as the time when these innovations occurred. Anthony (2007: 369; 379) suggests that the connection with Indo-Iranian implies that Pre-

Greek was spoken on the eastern border of southeastern Europe, perhaps as part of the early western Catacomb culture north of the Black Sea, 2800-2200 BCE (see Map 2).



Map 2: Culture groups of the Middle Bronze Age, 2800-2200 BCE (Anthony 2007: 379). (Reprinted by permission of Princeton University Press.)

North of this Catacomb cultural territory lies the Abashevo culture, 2500-1900 BCE, assumed by many Russian archeologists to be associated with Indo-Iranian speakers; the ceramics of this area and to some extent the metallurgy bear increasing similarity to those of the Catacomb culture, as steppe customs spread to the

forest-steppe in the north (2007: 382-3). Herders from this area, along with those of the Poltavka area, began settling further north and east, near the marshes of the Sintashta area in the northern steppe east of the Ural Mountains. Anthony (2007: 408-11) presents abundant evidence that the Sintashta settlement, dated between 2100 and 1800 BCE, was inhabited by predecessors of the Indo-Iranians: it contained the remains of the earliest chariots found anywhere, signs of metallurgical activity in each house, new kinds of weapons, and distinctive funeral rituals, all of which match descriptions in the Rigveda remarkably well.[17]

The predecessor of Iranian, according to Anthony, must have grown up in the northern steppes among Andronovo and Srubnaya people; the predecessor of Old Indic, on the other hand, must have developed further to the south, in the contact area of Central Asia (2007: 455). But the predecessors of Indo-Iranian itself must have been in contact with Pre-Greek at an earlier time.

3. Implications for the "Out of India" Theory

We can now turn briefly to an examination of how these considerations can help us assess the validity of the "Out of India" theory. As discussed above, the traditional explanation for the distribution of Indo-Aryan languages proposed by western scholars is that they entered the Indian subcontinent from the northwest sometime before 1500 BCE. Over the last century and a half, a strong reaction to this claim has been mounted by a number of Indian scholars, who state that Indo-Aryan is indigenous to India, and that no such migrations occurred. Some recent proponents of this interpretation can be mentioned:[18]

Eventually, from the foregoing discussion it can be inferred that the Sanskrit-speaking race, the Aryans, originally belonged to India. They emigrated to other parts of the globe to colonize alien lands and thus gave birth to the whole bulk of other Indo-European languages. This is why they still have a close affinity to their original source, Sanskrit. (Arya 1993: 57)

Let us wait for a 'co-operation' of various linguists from different branches. Let there be a constant confrontation of the linguistic and archaeological data. Let there be 'a co-herent analysis of the 'pre-IE' phenomenon'. Till then the Aryan problem remains unsolved and Aryan migration into India or an Aryan invasion of India remains in the fanciful dream of the Chatterjees and Ghoses. **We can at best entertain it as a myth, not a reality. The whole of Indo-European or Aryan theory [...] according to 'Science' is treacherous.** [Emphasis and punctuation as in original] (Choudhury 1993: 92)

Gupta (1993), likewise, expresses distrust for western views: he examines the theories which have been put forward concerning Brāhui's status as an isolated Dravidian language, whether it represents a remnant of earlier pervasive Dravidian presence in the north which was surrounded by an Indo-Aryan-speaking population, whether it was a group left behind as Dravidians entered India from the northwest, or whether it was a group which migrated up from the Deccan (cf. Emeneau 1962: 68-70). He concludes (1993:160) that Brāhui "*has nothing to do with the present-day South Indian Dravidian languages.*" [emphasis his] Claims like this, and his reluctance to see Dravidian as originally developing separately from Indo-Aryan, lead him to state that the Aryans were an "indigenous people [who] kept on living where they originally lived." (1993: 163)

Bryant (2001 and in this volume) conducts a thorough analysis of the claims made on both sides, remaining largely agnostic as to the validity of either set of arguments. One crucial claim that he repeats throughout the book is of special interest to us here: in his chapter "Linguistic evidence from outside of India," he concludes that the linguistic evidence, as represented by loanwords, does not exclude the Indigenous Aryanists' claim that Sanskrit arose on the Indian subcontinent (2001: 139). He later states the following in his conclusion:

Apart from the horse evidence, linguistic paleontology has not provided much uncontroversial data to exclude an eastern homeland that cannot be reversed to support the same; neither has the existence of loanwords, nor dialect geography, nor arguments based on homogeneity versus heterogeneity.

I would argue, on the contrary, that historical linguistics has provided clear evidence indicating that India cannot have been the indigenous homeland of Sanskrit or of PIE itself. Many arguments from a linguistic and an archeological standpoint have been put forward to counter the theory, as western and some Indian scholars engage in debate with its proponents. In an issue of the *Journal of Indo-European Studies* devoted especially to the Indo-Aryan Migration Debate (2002, Vol. 30, Numbers 3 & 4), for example, Kazanas (2002, and in this volume) presents a set of arguments in support of the Indigenous Aryan claim, and dismisses the validity of linguistics as a tool for assessing migration patterns or relationships. He states that "the issue of origins, of *when* and *how*, is one not for philologists but for archaeologists and experts in related fields" [emphasis his], to which Huld, in the same issue (2002: 356), responds that IE is a "linguistic concept which demands linguistic evidence" and that "unless one is willing to become conversant with linguistic theory, one cannot meaningfully participate in a discussion of Indo-European origins."

Zimmer, in the same issue (2002: 407), likewise uses linguistic arguments to respond, pointing out, among other things, the difficulty of explaining the remarkably archaic features of Old Irish, if one posits such a long move from an Indian *Urheimat*.

Mallory (2002), taking an archeological tack, proceeds to demonstrate the improbability of Kazanas' arguments by setting up potential models which could support Kazanas' Out-of-India theory, all of which fail completely. He points out that evidence is altogether lacking for westward movements from the Bactria-Margiana-Archaeological Complex (BMAC) northward and then westward into Europe, evidence which would be crucial to Kazanas' claim; rather, "[t]he antecedents of the steppe cultures east of the Urals are clearly west of the Urals where we find earlier evidence for domestic plants, animals, wheeled vehicles, metallurgy and the antecedent human physical type" (2002: 384).

A very lucid response to several of the claims of Misra (1992) is presented by Hock (1999), who also describes the type of scenario which would be necessary to support the "PIE in India" hypothesis. What would be needed, says Hock, is that migrations would have proceeded out of India, which preserved the dialectological relationships among IE languages as presented in Fig. 11.

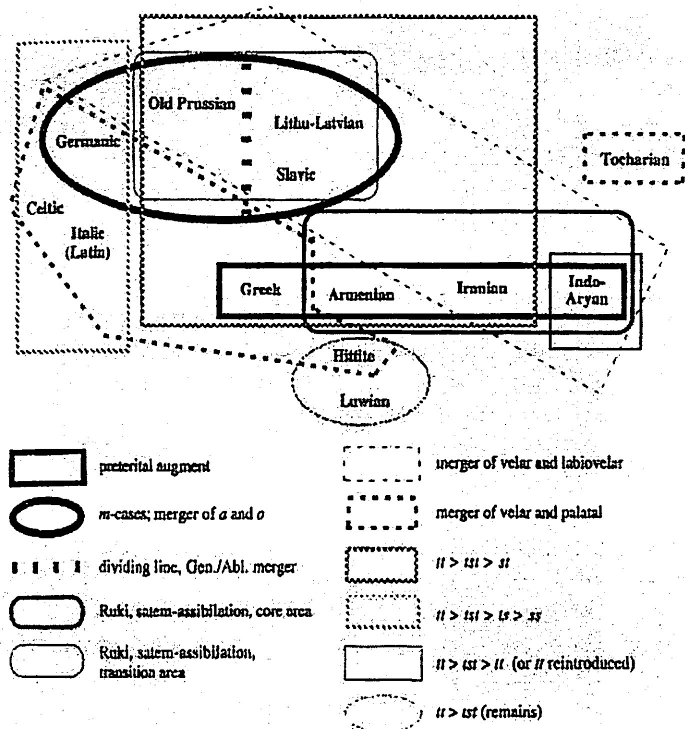


Fig. 11: A dialectological view of Indo-European (Hock 1999: 15). (Reprinted by permission of Harvard Oriental Series)

Hock points out that, if the simple family tree were the only model available, one could envision the exodus of groups in sequential manner, with the positioning of each of these groups, one by one, across the map. But such a model would not account for the similarities due to geographical contiguities noted in Fig. 11. Hock claims that if India were the *Urheimat*, then these relationships would have had to exist in India and would need to have been maintained as speakers passed through one of the "bottle-neck" routes out of India, a very unlikely occurrence.

Hock does not appear to envision another permutation of this scenario which Indigenous Aryanists might propose, that less differentiated IE languages could have emerged from India, and then have established the areal relationships noted in Figure 11. Proponents might argue that speakers moved out at different times and then came in contact with neighbors at a later time. It is this argument that I believe the issues discussed in this paper can best address. Not only should we avoid the simplification and distortion that the family tree model creates, as described by Hock, but we must also take note of another fact: a stratified model of PIE, one which takes into account the systematic, layered morphological correspondences of Greek and Indo-Iranian, cannot accommodate an Out-of-India explanation, for if we posit an Indian origin for Indic languages, how can we account both for the archaic morphological similarities which Sanskrit shares with many IE languages (e.g., a use of reduplication in the perfect) and those innovative features which it shares only with Iranian and Greek (e.g., a productive use of reduplication as an obligatory marker of the perfect)? We would need for Greek to have left India very late to account for these similarities, and to have made its way across a wide swath of terrain in very speedy and unified fashion. That is the difficult scenario which proponents of the Out-of-India approach would need to explain, if the traditional date of composition of the RV were maintained at c. 1500 B.C.E. If Kazanas' projection of 4500 B.C.E. for the writing of the RV is

adopted, we have no hope of accounting for these similarities or reconciling the evidence from stratification.

In the last analysis, we should note that a serious flaw in the Indic indigenous theory is that it presumes a family tree-type model: it posits that the "sister" languages occupy separate branches that did not interact. Bryant (2001: 300, and in this volume) claims that there is no linguistic evidence of loanwords or dialect geography to demonstrate a migration. But, if contact is given due consideration in historical analysis, we will note that Indo-Iranian and Greek must have been in contact at a fairly late time, if the morphological developments which we witness in both are to be accounted for. Should we position the Greeks in India to accommodate these correspondences? A much more likely scenario would be to view Pre-Indo-Iranian and Pre-Greek as neighbours in the steppes themselves, as Anthony (2007) does, and to regard them as late migrators. Contact linguistics thus provides crucial evidence for the migrationist hypothesis.

4. Conclusions

To sum up, the paper has presented the following three related arguments:

1) Family tree models are inadequate when used in isolation, and should be supplemented with more explanatory models which take contact into consideration. A three-dimensional model is proposed, which represents an amalgamation of family tree and wave models.

2) The implication of constructing such a model which takes "horizontal," areal influence into consideration is that the stratification of data—innovative layers building on more archaic layers—emerges as significant. Languages which share only archaic elements, such as Hittite and Germanic, are presumed to have separated from other IE languages at an earlier time; languages which share an array of morphological, lexical, poetic, and other features, like Indo-Iranian and Greek, are assumed to have remained in contact for a longer period.

3) The implication of this "late" contact of Indo-Iranian and Greek is that neither Indic not Indo-European itself can reasonably be assumed to have developed on the Indian subcontinent. Stratified morphological evidence, along with other convincing arguments, points to the steppe, not to India, as the most reasonable area of origination for PIE.

Endnotes

[1] Thanks to Joseph Roy for graphing the 3-dimensional tree, and to my graduate assistants, Kristina Gutierrez, Thomas Chu, and Michael Lee Gardin for their bibliographical and editorial help.

[2] See Blažek 2007 for a diachronic synopsis of IE tree models.

[3] Nakhleh et al. (2005) is more concerned with testing computational methods than with presenting data of interest to Indo-Europeanists, and so will not be analyzed in detail here.

[4] For the lexical analysis, the 208-word Swadesh meaning list is used, consisting of the lexical meanings deemed by Swadesh (1952) to be most basic and least prone to borrowing in the languages of the world. See, however, Kessler's convincing demonstration (2001:104-08) that a large number of the words on these lists have, indeed, been borrowed.

[5] Indic, Iranian, Slavic, Baltic, Armenian, and Albanian developed a spirant (*s* or *š*) from palatal **k'* where other IE languages have maintained a reflex of the ancient **k*, as found in the words for 'hundred': Skt. *šatám*, Av. *satəm*, Russ. *sto*, Lith. *šimtas* vs. Lat. *centum*, Gk. *hekatón*, Eng. *hundred* (Szemerényi 1990: 59-61).

[6] Indic, Iranian, Slavic, and, to some extent, Baltic changed *s* to *š* after *r, u, k*, and *i*, appearing as retroflex ṣ in Sanskrit and *x* in Slavic before back vowels, e.g.: **ters-* 'dry' Gk. *térsomai* 'I become thirsty', Lat. *tostus* (< **torsitos*), Eng. *thirst* vs. Skt. *tṛṣyati* 'thirsts', Av. *taršna-* 'thirst', Lith. *tirštas*. Likewise, **wers-* 'high place' Lat. *uerrūca* 'wart' vs. Skt. *varṣman-* 'height, peak', Lith.

viršŭs 'summit', OCS *vrĭxŭ* 'summit, height' (Szemerényi 1990: 51-2). (See Andersen (1968) and in this volume for a detailed analysis of Baltic and Slavic; see also Drinka (1995a: 35-37) for discussion and references.)

[7]In later work (e.g., Ringe et al. 2002: 109), the genetic nature of the "satem core" is downplayed, and the role of contact is, instead, acknowledged, following Hock (1986: 442-44). But these innovations are still used as data to help establish genetic connection.

[8] Taylor et al. stress that their analysis does not constitute an endorsement of the Italo-Celtic hypothesis, however (2000: 404).

[9] Cf., for example, the numerous isoglosses for Celtic and Germanic (Porzig 1954: 118-27) and for Germanic and Balto-Slavic (140-48), as well as Porzig's discussion of the Italic, Germanic, Baltic, and Slavic contact zones (205-09), and his concluding chapter on the distribution of IE dialectal features (213-17).

[10]"To be sure, the effects of imperfect second-language learning strongly resemble 'borrowing of morpho-syntax' after the fact, if the imperfectly learned second language eventually becomes a community norm. But in our view such a pattern in the data reveals a discontinuity of transmission which should exclude the language in question from any strict 'family tree.'" (Ringe et al. 2002: 107)

[11]McMahon (1994: 248) calls this issue the "real actuation question": "why some of these innovations die out and others catch on, spreading through the community, or why certain instances of variation become changes while others don't."

[12]They also note, to their credit, that nodes are to be regarded as capable of containing variation (Ringe et al. 2002: 108).

[13] When faced with the challenging distributions such as that of Germanic, Ringe et al. propose that other models, such as network models of linguistic diversification, might be more suitable for such difficult cases (Ringe et al. 2002: 108-112). See section 1.4 for further discussion.

[14] McMahon and McMahon (2005: 109) constructed this least conservative sub-list by choosing words from the list of Dyen et al. (1992) which had "less potential for cross-cultural generalizability" (e.g., 'grass', 'year', 'think'), and which were thus expected to be more subject to the influence of contact.

[15] Watkins (2001: 57) notes that, besides sharing the augment prefix and the "whole structural organization of the verbal system," Indo-Iranian and Greek also have more poetic features in common than any of the other IE languages. Watkins places Indo-Iranian, Greek, Armenian, and to some extent Phrygian in a large subgroup based on these features, and draws a schematic branching diagram with these languages on the right edge. He admits, however, that other models, including those based on geography, are also possible. The judgment of Kessler (2001: 195) is also germane here: he suggests that grammatical morphemes are more difficult to use in the statistical analysis of relatedness than word lists, since the former are harder to match than the latter, and are subject to analogical reshaping. One could comment that, when strong resemblances actually do appear, one should take particular note.

[16] Besides morphological evidence, Anthony (2007: 55-56) also collects important cultural information, such as shared words for weaponry, ritual sacrifice, deities, and poetic meter.

[17] For example, besides a large number of horse sacrifices at this site, evidence of midwinter dog sacrifices were also found at nearby sites, suggesting the New Year's initiation sacrifices referred to in the RV (Anthony 2007: 410).

[18] Certainly not all Indian scholars are proponents of Aryan indegeneity; Shereen Ratnagar (1999), along with others, presents data and cogent arguments in favor of movement into India from without.

References

Aikhenvald, Alexandra, and R. M. W. Dixon (eds.)
2001 *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics.* Oxford: Oxford University Press.

Andersen, Henning
1968 IE *s after *i, u, r, k* in Baltic and Slavic. *Acta linguistica hafniensia* 11: 171-90.

Andersen, Henning (ed.)
1995 *Historical Linguistics 1993.* Amsterdam/ Philadelphia: Benjamins.
2003 *Language Contacts in Prehistory: Studies in Stratigraphy.* Amsterdam: Benjamins.

Anthony, David
2007 *The Horse, the Wheel, and Language.* Princeton / Oxford: Princeton University Press.

Arya, R. P.
1993 The Aryan problem: A linguistic approach. In: Dao and Kamath (eds.). 51-57.

Birwé, Robert
1955 *Griechisch-arische Sprachbeziehungen im Verbalsystem.* Waldorf-Hessen: Verlag für Orientkunde.

Blažek, Václav
2007 From August Schleicher to Sergei Starostin: On the development of the tree-diagram models of the Indo-European languages. *Journal of Indo-European Studies* 35: 82-109.

Bronkhorst, Johannes, and Madhav Deshpande (eds.)
1999 *Aryan and Non-Aryan in South Asia: Evidence, Interpretation and Ideology.* Cambridge, MA: Harvard.
Bryant, Edwin
2001 *The Quest for the Origins of Vedic culture: The Indo-Aryan Migration Debate.* Oxford: Oxford University Press.
Bybee, Joan, Revere Perkins, and William Pagliuca
1994 *The Evolution of Grammar.* Chicago: University of Chicago Press.
Campbell, Lyle
2006 Areal linguistics: A closer scrutiny. In: Matras et al. (eds.) 1-31.
Chappell, Hilary
2001 Language contact and areal diffusion in Sinitic languages. In: Aikhenvald and Dixon (eds.) 328-57.
Choudhury, Paramesh
1993 *The Aryans. A Modern Myth.* Part I. New Delhi: S. Chakraborty.
Dao, S.B., and Suryanath Kamath (eds.)
1993 *The Aryan Problem.* Pune: Bharatiya Itihasa.
Dixon, R. M. W.
1997 *The Rise and Fall of Languages.* Cambridge: Cambridge University Press.
Drinka, Bridget.
1993 The dispersion of Indo-European dialects: Clues from morphology. *Word* 44: 413-29
1995a *The Sigmatic Aorist in Indo-European: Evidence for the Space-Time Hypothesis.* Monograph 13 in the Journal of Indo-European Studies Monograph Series.
1995b Areal linguistics in prehistory: Evidence from Indo-European aspect. In: Andersen (ed.) 1995. 143-58.
1999 Alignment in Early Proto-Indo-European. In: Justus and Polomé (eds.) 462-98.
2003a Areal factors in the development of the European periphrastic perfect. *Word* 54.1: 1-38.
2003b The development of the perfect in Indo-European: Stratigraphic evidence for prehistoric areal influence. In: Andersen (ed.) 2003. 77-105.
Dyen, Isidore, Joseph Kruskal, and Paul Black
1992 An Indoeuropean classification: a lexicostatistical experiment. *Transactions of the American Philosophical Society* 82/5: 1-132.
Emeneau, Murray B.
1962 *Brahui and Dravidian Comparative Grammar.* Berkeley / Los Angeles: University of California Press.
Euler, Wolfram
1979 *Indoiranisch-griechische Gemeinsamkeiten der Nominalbildung und deren indogermanische Grundlagen.* Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Garrett, Andrew
1999 A new model of Indo-European sub-grouping and dispersal. Berkeley Linguistics Society 25: 146-56.
2005 Convergence in the formation of Indo-European subgroups: Phylogeny and chronology. In: Forster and Renfrew (eds.) 139-51.
Forster, Peter and Colin Renfrew
2005 *Phylogenetic Methods and the Prehistory of Languages.* Cambridge: McDonald Institute for Archaeological Research.
Gupta, S. P.
1993 Palaeo-anthropology and archaeology of the Vedic Aryans. In: Deo and Karnath (eds.). 153-65.
Harrison, S. P.
2003 On the limits of the Comparative Method. In: Joseph and Janda (eds.). 213-43.
Heine, Bernd, and Tania Kuteva
2005 *Language Contact and Grammatical Change.* (Cambridge Approaches to Language Contact) Cambridge: Cambridge University Press.
2006 *The Changing Languages of Europe.* Oxford: Oxford University Press.
Hock, Hans Heinrich
1986 *Principles of Historical Linguistics.* Berlin: Mouton de Gruyter.
1999 Out of India? The linguistic evidence. In: Bronkhorst and Deshpande (eds.) 1-18.
Huld, Martin
2002 Linguistic science, truth, and the Indocentric Hypothesis. *Journal of Indo-European Studies* 30: 353 – 63.
Jankowsky, Kurt R. and Ernst S. Dick (eds.)
1982 *Festschrift für Karl Schneider.* Amsterdam/ Philadelphia: Benjamins.
Joseph, Brian and Richard Janda
2003 *The Handbook of Historical Linguistics.* Malden, MA / Oxford / Melbourne / Berlin: Blackwell.
Justus, Carol F. and Edgar C. Polomé (eds.).
1999 *Language Change and Typological Variation: Papers in Honor of Winfred P. Lehmann on the Occasion of his 82nd Birthday.* Vol. II: Grammatical Universals and Typology. Monograph 31 in the Journal of Indo-European Studies Monograph Series.
Kazanas, N.
2002 Indigenous Indo-Aryans and the Rigveda. *Journal of Indo-European Studies* 30: 275-334.
Kessler, Brett
2001 *The Significance of Word Lists.* Stanford: CSLI Publications.
Mallory, James P.
2002 Indigenous Indo-Aryans: The Preservation and Total Distribution Principles. *Journal of Indo-European Studies* 30: 375-87.
Marcantonio, A.

2002 *The Uralic Language Family: Facts, Myths and Statistics.* Oxford/Boston: Blackwell
Matras, Yaron, April McMahon, and Nigel Vincent
2006 *Linguistic Areas: Convergence in Historical and Typological Perspective.* Basingstoke, Hampshire / New York:Palgrave Macmillan.
McMahon, April
1994 *Understanding Language Change.* Cambridge: Cambridge University Press.
McMahon, April and Robert McMahon
2005 *Language Classification by Numbers.* Oxford: Oxford University Press.
2006. Keeping contact in the family: Approaches to language classification and contact-induced change. In: Matras et al. (eds.) 51-74.
Meid, Wolfgang
1975 Probleme der räumlichen und zeitlichen Gleiderung des Indogermanischen. In: Rix (ed.) 204-19.
Meillet, Antoine
1931 Essai de chronologie des langues indo-européennes. *Bulletin de la Société de Linguistique de Paris* 32: 1-28.
Misra, Satya Swarup
1992 *The Aryan Problem: A Linguistic Approach.* New Delhi: Munshiram Manoharlal.
Mufwene, Salikoko
2001 *The Ecology of Language Evolution.* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press.
Nakhleh, Luay, Tandy Warnow, Don Ringe, and Steven N. Evans
2005 A comparison of phylogenetic reconstruction methods on an Indo-European dataset. *Transactions of the Philological Society* 103: 171-92.
Polomé, Edgar C.
1964 Diachronic development of structural patterns in the Germanic conjugation system. *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., 1962.* The Hague: Mouton. 870-80.
1982 Germanic as an archaic Indo-European language. In: Jankowsky and Dick (eds.). 51-9.
1987 Who are the Germanic people? In Skomal and Polomé (eds.) 216-44.
Porzig, Walter
1954 *Die Gliederung des indogermanischen Sprachgebiets.* Heidelberg: Winter.
Ratnagar, Shereen
1999 Does archæology hold the answers? In: Bronkhorst and Deshpande (eds.) 207- 38.
Ringe, Don, Tandy Warlow, and Ann Taylor
2002 Indo-European and computational cladistics. *Transactions of the Philological Society* 100: 59-129.
Rix, Helmut (ed.)

1975 *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, September 1973.* Wiesbaden: Reichert.
Schleicher, August
1860 *Die Deutsche Sprache.* Stuttgart: Cotta.
Schmidt, Johannes
1872 *Die Verwandtschaftverhältnisse der indogermanischen Sprachen.* Weimar: Böhlau.
Skomal, Susan Nacev, and Edgar C. Polomé (eds.).
1987 *Proto-Indo-European: The Archaeology of a Linguistic Problem.* Washington, D.C.: Institute for the Study of Man.
Smith, John Charles, and Delia Bentley (eds.).
2000 *Historical linguistics 1995.* Vol. I: General Issues and Non-Germanic Languages. Amsterdam / Philadelphia: Benjamins.
Swadesh, Morris
1952 Lexico-statistic dating of prehistoric ethnic contacts. In: *Proceedings of the American Philosophical Society* 96: 453-63.
Szemerényi, Oswald
1996 *Introduction to Indo-European Linguistics.* (Translated from *Einführung in die vergleichende Sprachwissenschaft.* 4th ed., 1990, with additional notes and references). Oxford: Oxford University Press.
Taylor, Ann, Tandy Warlow, and Don Ringe
2000 Character-based reconstruction of a linguistic cladogram. In: Smith and Bentley (eds.). 393-408.
Thomason, Sarah
2001 *Language Contact.* Edinburgh: Edinburgh University Press.
Thomason, Sarah G., and Terrence Kaufman.
1988. *Language Contact, Creolization, and Genetic Linguistics.* Berkeley / Los Angeles / London: University of California Press.
Trudgill, Peter.
1983 *On Dialect: Social and Geographical Perspectives.* NewYork /London: New York University Press.
Watkins, Calvert
2001 An Indo-European Linguistic Area and its characteristics: Ancient Anatolia. Areal diffusion as a challenge to the Comparative Method? In: Aikhenvald and Dixon (eds.) 44-63.
Wolfram, Walt, and Natalie Schilling-Estes
2003 Dialectology and linguistic diffusion. In: Joseph and Janda (eds.) 713-35.
Zimmer, Stefan
2002 The easy way to truth: "Heaven does not lie." *Journal of Indo-European Studies* 30: 401-09.

The Origin and Spread of the Indo-Germanic People

Alexander Häusler
Halle (Saale, Germany)

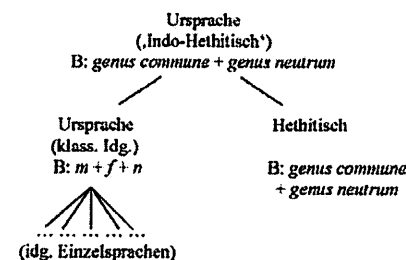
There are no archaeological proofs of the existence of an ancient *Urkultur* of the Indo-Europeans, with a specific social system and religion, all confined within a well defined, restricted *Urheimat*. Nevertheless, many different explanations for the origin and spread of the IE languages, peoples and culture have been proposed in the course of the years. The Author reviews these models of interpretation, including the 'migrationist' and 'diffusionist' ones, and concludes that there is certainly no evidence in favour of the migrationist models, whilst the diffusionist theses do not appear to contradict the archaeological evidence. Even so, the (assumed) IE linguistic community is not at all identifiable by the means of archaeological or anthropological research, although it must have existed, since its corresponding linguistic classification has been fairly safely established.

1. Introduction: Indo-Germanic vs Indo-European

Two synonyms are used to refer to the language family under discussion: Indo-Germanic and Indo-European. Here I shall use the term Indo-Germanic (Idg.) because this has been the first term used historically. 'Indogermanische Sprachen' and 'Indogermanen' are purely linguistic terms. The well-known linguist Trubetzkoy (1968: 214, originally 1939) wrote: "Indo-Germanics are such people, whose mother tongue belongs to the Indo-Germanic language family. From this, which is the only possible academic definition, it follows that 'Indo-German' is a purely linguistic term, like 'syntax', 'genitive', 'consonant shift' etc. There are Indo-Germanic languages and there are peoples who speak these languages. The only thing which all these peoples have in common is the affiliation of their languages to the same

language family". In the same way, Feist (1926: 54) states that: "The Indo-Germanics ... are not a historically identifiable people, but only an abstraction from linguistic facts".

One of the major tasks of Idg. studies consists in reconstructing an Idg. (/ proto-Idg.) common, parent language from languages distributed over large parts of Eurasia - from island Celtic in the west to Tocharian in the Tarim basin in Central Asia (Fig. 1).



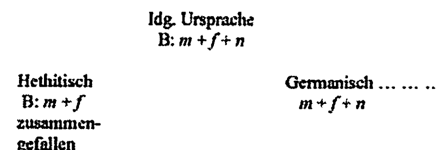


Fig.1. Classification of the Indo-Germanic languages: a) including Indo-Hittite, b) Classic model, without Indo-Hittite.

All the languages involved show a high number of correspondences in all possible areas, which are firmly linked together and are not in any way to be explained by chance. The traditional explanation for the numerous correspondences in the

fields of phonetics, morphology and semantics is seen in the assumption of a common origin and the subsequent process of diffusion (Meier-Brügger 2002: 57 ff.), a process typically represented through the well know family tree diagram (Fig. 2).

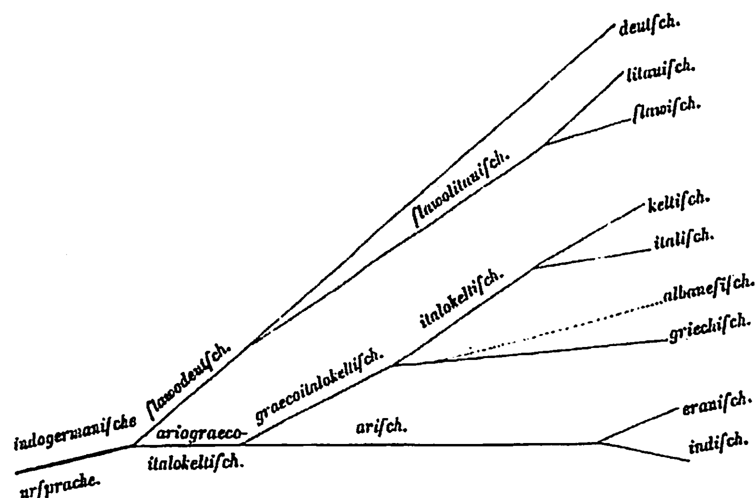


Fig. 2. Tree models of the origins of Idg. languages, after A. Schleicher 1861.

This being the case, many models have been proposed through the years to explain when, how and why the Idg. languages spread across such a wide Eurasiatic area.

Before delving into this topic, let us remember that, in contrast to what happens in archaeology, within the field of linguistics, much later periods, much later phases of development of the languages in question are taken into consideration (Schlerath 1992: 137), because of the very nature of the linguistic investigation. Typically, linguists operate on a time scale that spans only a few centuries before the attestation of the oldest Idg., single languages, that is, approximately the mid-3rd millennium B.C. (Schmitt 2000: 395).

1. The computer simulation of J. Robb

According to Robb (1991, 1993), under conditions of an increased mobility of sections of the population (Palaeolithic, Mesolithic), and assuming statistical, purely coincidental displacements of languages over millennia (from an originally very large number of languages), it turns out that some languages may be displaced and die out, while others extend further, and continually, at the expense of the lost languages. This model predicts that only a few language families, from an original language plurality, are finally left in a territory, in this case: Eurasia. After an initial phase of language displacement (in the Palaeolithic), there could also be phases of language consolidation during a period of increased 'sedentariness'. In his computer simulation Robb assumes a time zero (Fig. 3a) as a starting point of this process. In the distribution map every territory is labelled with an alphabetic character indicating a given language family. In this way the map of a hypothetical language distribution at time zero comes into being. The hypothetical factor of the displacing of a language by an adjacent one is set at 35%. After 30 cycles (turns) 16 of 64 language families prove to be extinct, after 60 cycles only 28 are left (Fig. 3b), and after 120 cycles only 20 of 64 language families are left (Fig. 3c). The major language families then cover a spatial extent which corresponds to that of the Idg. languages towards the end of prehistoric times. After 450 cycles only two language families are left (Fig. 3d). From this model one could draw the conclusion that the specific language system of Idg. does indeed go back to one genuine creation, one nucleus (extreme formulation: at a campfire). According to Robb the cradle of Idg. might have lain in a very restricted territory, and its origin may

date back to very ancient times, that is, to a period much further back in time than conventionally assumed, though, this model is unable to specify a definite date for the *Ursprung* of the Idg. language system.

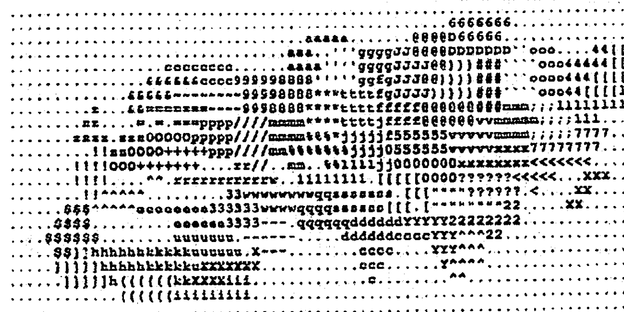


Fig. 3a. Computer simulation of the spread of languages towards a factitious point of origin.



Fig. 3b. Computer simulation of the spread of languages after 60 cycles.



Fig. 3c. Computer simulation of the spread of languages after 340 cycles.

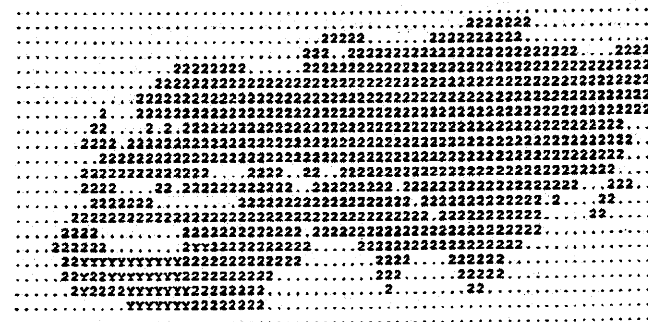


Fig. 3d. Computer simulation of the spread of languages after 2000 cycles.

In this context, one should bear in mind that a much earlier time period extending to the 7th / 6th millennium B.C. has been considered when modelling the origin and spread of the Finno-Ugric language, According to the results of Finnish archaeology, the ancestors of the Finns (the Finno-Ugric peoples) were already resident in Finland and the adjacent regions for many millennia, this thesis being in turn in contradiction with the prediction made by the conventional Uralic theory, that the Finns have migrated into their present settlement area only since 500 B.C. (Nuňez 1998, 2000, 2002). The much discussed contacts between the Finno-Ugric peoples / languages and the more southerly and westerly disseminated Idg. could have taken place therefore,

possibly, already several millennia earlier than conventionally assumed.

2. The emergence of Indo-Germanic through language fusion

In the past, numerous renowned linguists often explained the emergence of languages, including Idg., as a merging together of many smaller language groups in the course of time into a more comprehensive language area, in this case the area of the Idg. parent language (Trubetzkoy 1968: 214; Scherer 1968: 303). Today, the thesis of the melting together of many smaller linguistic units into one common language structure is not that widely accepted, because such a process is not verifiable in historical times. Nevertheless, the convergence model is still at times adopted. For example, Dixon (1997) explains the emergence of the Pama-Nyungan languages of Australia by such a process of convergence, the mixing of about 100 different languages within one diffusion area. Renfrew (1999, 2003) on the other hand is inclined to explain the emergence of Idg. during the Neolithic phase of Anatolia, about 9,000 years ago, through the very same process of language mixing (see below).

3. Schmid's model: The 'Old Europe' concept

In the following, the 'Old Europe' concept, as developed by Schmid and Udolph, shall be reported. The Old Europe concept is based on the examination of the old names of water bodies in Europe, *Old European hydronymy* (Udolph 1994; Schmid 1985, 1987). This concept assumes that a morphologically and etymologically uniform naming of bodies of water must have once existed in extensive and mostly joined areas of Europe, of which significant remains are still in use to this day. At this point it must be noticed that the archaeological term 'Old Europe' after M.

Gimbutas, and the 'Old Europe' concept proposed by Udolph and Schmid are fundamentally different concepts. For Schmid, as mentioned, the term *Alteuropa* refers to the names of certain bodies of water (rather than to the languages), and these names in turn can be assigned to certain, geographically well defined, places or bodies of water. These Old European names presume the totality of the Idg. vocabulary, including the vocabulary of those languages which themselves are not involved in this hydronymy, like Indo-Iranian. "Fügt man nun die alteuropäischen Sprachen zu einer abgrenzbaren Namenlandschaft zusammen, so sind sie Ausdruck eines sprachlichen Kontinuums, das von Spanien bis an den Don, von Skandinavien bis in die Mittelmeerländer reicht" (Schmid 1983: 111). "Dabei ist auf dem Territorium der alteuropäischen Hydronymie keine ältere Sprachschicht erkennbar, die älter als ein idg. Sprachkontinuum wäre" (Schmid 1983a: 404). "Das postulierte Alteuropäisch ist weder jünger noch verschieden vom rekonstruierten Idg., sondern mit ihm identisch" (Schmid 1968: 253). It is important to observe here that, according to Schmid (1968: 247), the uniform Old European hydronymy also presupposes a uniform linguistic community, although the frequently held view that Old European represents an interlayer between Idg. and the various, separate Idg. languages is not correct [Schmid 1995; for previous criticism of the Old Europe conception see Schmid (1998)]. The *Alteuropa* concept presupposes that in an area extending at least from Scandinavia to the Don a language *continuum* must have existed that can already be described as Idg. Thus, the transformation of this language *continuum* into the single languages / language branches (such as Greek, Celtic, Germanic, Baltic etc.) must have taken place in the respective regions, this language *continuum* being itself already regionally differentiated. A temporal priority befits none of the sub-areas. In other words, Schmid identifies a large territory in Europe with the Old

European, IE languages. This model does not contradict the thesis of very old connections between IE and Finno-Ugric (see below).

The Old Europe concept, despite not being widely accepted – since the Old European hydronymy is not distributed evenly in Europe – does not contradict the fusion hypothesis proposed by Trubetzkoy. Furthermore, it harmonizes with the results of my research, according to which one can assume a continuous, progressive development of the autochthonous populations and their culture, undisturbed by migration and conquest, since the Mesolithic age (at the latest) for large parts of Europe (including the settlement areas of the Celts, Germans, Balts, Slavs and Greeks, as well as the north Pontic area, where the Indo-Iranian language speaking Scythians and Sarmathians lived). These conclusions are in turn in agreement with the results obtained by Finnish scientists, according to whom in Finland too one should assume a continuous, progressive development of the old-established population – together with their culture and language – uninfluenced by migrations since the Mesolithic period.

4. Indo-Germanic – Finno-Ugric language contacts

With reference to the Finno-Ugric languages, and, more generally, to the Uralic region, the renowned Finnish scholar, J. Koivulehto (2003:309), reaches the general conclusion that: "Die älteste engere Urheimat – wenn nicht der Ururalier, dann wenigstens die der westlicheren Finno-Ugrier – sich etwa von der Ostsee bis an den Ober- und Mittellauf der Wolga erstreckt hat". With regard to the thesis of the assumed contacts between the Idg. and the Finno-Ugric populations, the Author states the following (2003: 310): "Die 'gemeinsamen' Wörter lassen sich lautlich einwandfrei als frühe Entlehnungen mit angemessenen lautlichen Substitutionen erklären. Bei Urverwandtschaft wäre größere lautliche Entfernung zu erwarten [....] Aus der zentralen Natur des entlehnten

Wortmaterials, oft verbaler Natur, ... muss man folgern, dass die Kontakte sehr intensiv gewesen sind". The (assumed) contacts between the Idg. and the Finno-Ugric languages must have been confined to the western regions, as testified by the nature of the loan words. In fact, "An den Lehnwörtern läßt sich ablesen, dass es im nördlichen Ostseeraum, in Skandinavien, Finnland und im Baltikum eine indogermanisch-vorgermanisch-germanische sprachliche Kontinuität ohne erfassbare Lücken gegeben hat" (Koivulehto 2003:311). According to Koivulehto, the Idg. / Finno-Ugric (/ Uralic) contacts existed since the primary language phase of both language families, the oldest loan words being common across the whole of today's Uralic language area. On the other hand, there would be no indication of an ancient kinship.

The thesis outlined above about the contacts between the ancient Finno-Ugric and Idg. languages is very much in harmony with the hypothesis of an 'Old European', Idg. language *continuum*, perhaps existing already since the Mesolithic age in the area extending between the North Sea and the Caspian Sea.

5. Schlerath's model

5. 1. I shall present here another model to explain the origin of the Idg., the one proposed by Schlerath, that mediates between migration and autochthony. Schlerath (1973: 24) states that: "Das von uns rekonstruierte Protoidg. ist die Sprache, die sich bei den wandernden Kriegern herausgebildet hat. Die Protoidg. erhielten laufend Zuzug vor allem von der Jungmannschaft der seßhaften Bevölkerung. ... Wenn das Protoidg. die Sprache des erobernden Kriegeradels war, erklärt sich auch seine verhältnismäßig große Einheitlichkeit". Schlerath, also relying on the results obtained by the ethnologist W. E. Mühlmann (1968), assumes warlike invasions by relatively small clusters of peoples which were oddly assorted ethnically, but very uniform linguistically. In his opinion, in the linguistic material of the reconstructed, (uniform) proto-Idg.

two historico-cultural elements must have been present: a) the culture of that folk out of which those warriors who had asserted themselves linguistically (within the colourful, motley ethnic mix of the conquest period) emerged (this folk could have been a small or large group, and it must have engaged in agriculture and livestock breeding); b) the culture of those warriors who were the real bearers of 'Indo-Germanization'. This means that the reconstructable Idg. lexical material is based on the contribution of both cultures (Schlerath 1987:156). In an earlier work Schlerath (1981: 198) states that migrations of peoples practicing agriculture over large distances could only have taken place under the umbrella of military operations. These military campaigns, in the course of their wide ranging operations, must have been gaining an inflow of young male participants from other, unrelated peoples. Thus, the warlike operations must have been much more important for the expansion of the Idg. than the movements of populations farming under this military umbrella, or following these military campaigns. In other words, cultural cohesion was given by the warrior ideology. These considerations are clearly based on the thesis that there always existed a symbiosis between settled farmers and nomads: the nomadic stock-breeders were sifted out from settled farming people with whom, however, the nomads remained in constant contact, through diverse exchange relations, including, first of all, economic relations (Schlerath 1987: 256; Mühlmann 1968). Thus, Schlerath explains the origin and distribution of the Idg. languages by assuming the spread of some ruling groups, without mentioning a concrete, initial territory nor a specific span of time. However, Schlerath's basic assumptions do not appear to be valid for those millennia under consideration, because we cannot find any archaeological evidence supporting them.

5. 2. Thousands of specific archaeological complexes are known from the conjectured expansion areas of the Idg. in the steppe regions of eastern Europe, the north Caucasus and the eastern conjoining areas of north Eurasia and Central Asia, respectively. These findings date back from the Mesolithic phase up to the 3rd millennium B.C., and derive from hunters and gatherers, farming populations and those human groups for whom animal husbandry played a more or less significant role. The number of archaeological complexes known from central and northern Europe west of the Weichsel runs into hundreds of thousands, but here nothing indicates a central Asian, eastern European, north Pontic, north Caucasian nor Transcaucasian origin. Furthermore, no graves of isolated groups of warriors exist. What really can be proven is the existence of the burial mounds concentrated in the Hungarian lowland plain east of the Theiss, with the characteristic burials of the Ochre Grave Culture of the 3rd millennium B.C. (Fig. 4), which might actually be connected with the east European steppe regions in the area between the Volga and Dniester. However, nothing of significance was left by this population, which appears to have vanished without trace (Häusler 1998). As to Europe, the spread of farming, starting from Anatolia, can be traced from the 7th millennium B.C. In the European area north of the Alps the practice of farming has been proven to exist at least since the time of the first piece of evidence for the Linear Pottery Culture, ca. 5600 B.C. (Lüning 2000). Different interpretations have been proposed concerning the sociological conditions within the farming based Linear Pottery Culture. There is, however, no archaeological evidence suggesting either a military elite, or groups of warriors which could have come from different cultural areas, or a warrior ideology.

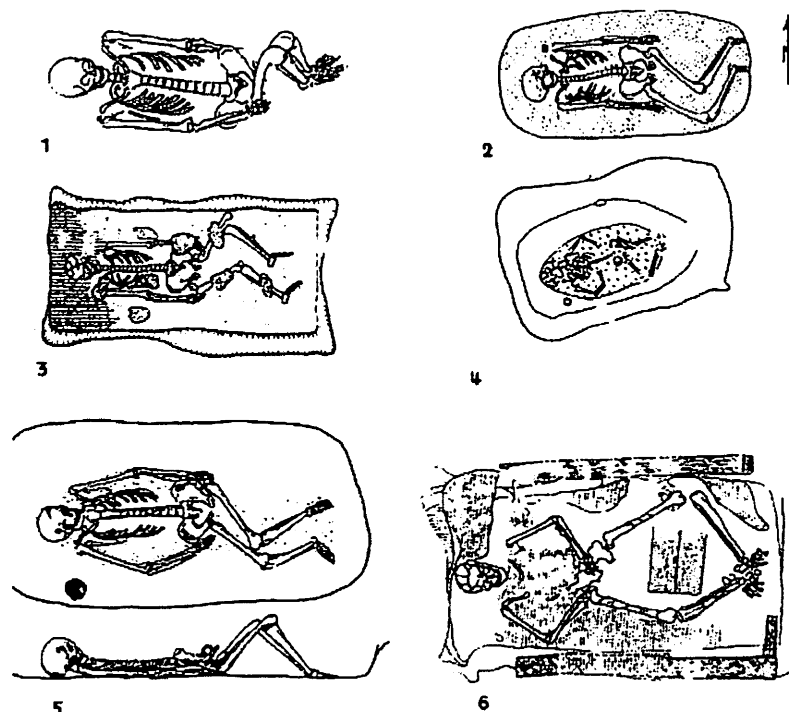


Fig. 4. Crouched burial of a western off-shoot of the Ochre Grave Culture in West Hungary. Skeleton oriented to the west, facing east.

In the post Linear Pottery time of the 4th millennium B.C., a continuous, progressive development of the old-established sections of the population and their cultures can be detected in southeast and central Europe. The same took place in the wide areas of northern Europe in the 3rd millennium B.C. (areas that are located before the regions of the origin of the oldest farming in the Near East). Similarly, in Europe, between southern Greece and southern Scandinavia, over the course of millennia, a continuous replacement of an economy based on hunting, gathering and

fishing by one based on farming and animal husbandry took place (Lüning 2000). Schlerath's 'emerging' warriors (Fig. 5) must have come from agricultural areas that had existed in northern Eurasia since millennia, and would have owned a firm language structure. This language structure would have been adopted completely by those young male members who joined the military campaigns, if they ever existed. This, of course, cannot be proven linguistically, and archaeology has largely dissociated itself from the formerly widespread concept of *Mischkultur*. Furthermore, it has to be taken into account that the existence of nomads in northern Eurasia is only ascertained since the middle or end of the 2nd millennium B.C. (Shilov 1989; Khazanov 1994).



Fig. 5. Idea of an Indo-Germanic mounted warrior of the Corded Ware Culture.

Thus, if there were no nomads yet – at least in the area between the Rhine, the Urals and Central Asia – within the early millennia under discussion, the thesis of their coexistence with the old farming societies can easily be discarded. The spread of farming in

Europe took place during the Neolithic in a peaceful way, without the help of warlike conquests, and that process stretched over millennia. Thus, the existence of warrior groups which lived outside the closed settlement area of farming cultures remains highly speculative, and, in fact, it cannot be ascertained archaeologically either. Only the still 'non-neolithicized' hunters, gatherers and fishermen existed there at that time.

6. Was there an Idg. *Urkultur*, in a specific *Urheimat*?

6. 1. We now come to a topic that has always attracted much attention: the issue of a specific, ancient homeland of the Idg., as well as their culture and way of life. Thus far, there have been many attempts to explore the original culture, lifestyle, social system and religion of the Idg. community (supposedly located in a specific, restricted *Urheimat*) with the help of the so called linguistic palaeontology. This *Urheimat*, in turn, was often brought into connection with certain archaeological cultures (Fig. 6).

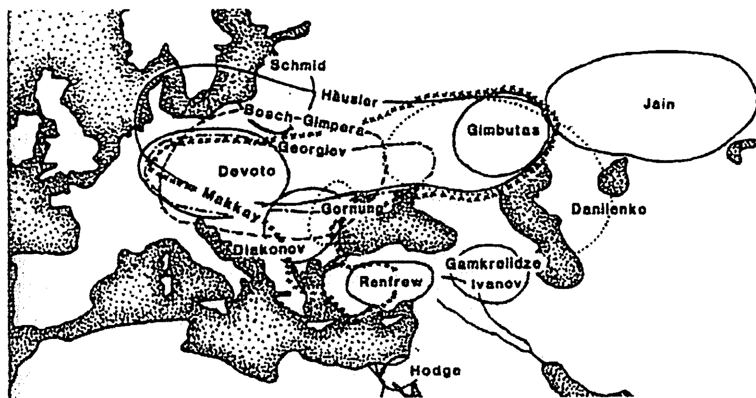


Fig. 6. Proposals to locate Idg. homelands.

The current state of research on the assumed culture and way of life of the Idg., is recorded both in lexica (Mallory & Adams 1997; see Häusler 2000) and in current overviews and textbooks (see for example Schmitt & Häusler 2000; Zimmer 2002, 2004). There is almost no major publication on the topic of the Idg. that would not take these problems into account, more or less extensively, despite the many warnings against reaching hasty and misleading conclusions at this regard. For example, about the rash attempts to reconstruct the culture, way of life, and migratory routes of the Idg, Pulgram (1968: 466) wrote that: "In manchen zeitgenössischen Kreisen solche Naivität als besonderer Scharfsinn zu gelten scheint, wie jeder sehen kann, der einige von den zahlreichen Bänden liest, die von den 'Indoeuropäern', von ihrem Leben und ihren Sitten handeln. Und wenn es nicht erstaunlich genug ist, daß solche Werke überhaupt verfaßt werden, so muß man sich doch jedenfalls wundern über die unkritische und bewundernde Gläubigkeit, mit der sie von vielen Forschern aufgenommen werden". On this issue Feist (1926: 36) states the following: "Aus dem für die Ursprache durch den Vergleich der verschiedenen idg. Sprachen ermittelten Wortvorrat gewinnt man nur unbestimmte Andeutungen für die Lokalisierung der Urheimat" (on his conclusion little has changed to this day). Similarly Untermann (1985: 161) claims: "Das Bild, daß sich hier abzuzeichnen beginnt - Haus, Dorf, Acker, Wald, Unwegsamkeit, Fremde - paßt nicht recht zu einer allseits offenen Steppenlandschaft mit ungehindert umherschweifenden Nomadenscharen. Es paßt besser zu einer Bevölkerung, die sich durch Rodung und Kultivierung ihre kleinen festumgrenzten Lebensräume schafft". In reality, the existence of a 'culture' of the Idg., within a specific *Urheimat*, is no more than fiction; therefore, seeking a connection between Idg. and certain anthropological types or races is, already in itself, methodologically wrong. As Trubetzkoy (1968: 216) puts it: "Man diskutiert über den

Wohnsitz, die Rasse und Kultur des angeblichen indogermanischen Urvolkes, welches vielleicht nie existiert hat... Man läuft einem romantischen Hirngespinste nach, statt sich an die einzige positive wissenschaftliche Angabe zu halten - nämlich, daß 'Indogermane' ein rein linguistischer Begriff ist" [this applies particularly to the reconstruction of a social structure within a proto-Idg. culture, as pointed out, among others, by Schlerath (1987: 256) and Zimmer (2004:7)]. Within this context one should mention the '*Idéologie tripartie*' propagated by Dumézil (1958), according to which the Idg. would already have had a trinomial social system consisting of priests, warriors and farmers. However, this theory is not supported by linguistic arguments nor any etymological comparisons. No proof of correctness is provided for the existence of this social structure reflecting on older (or often also later) textual attestations in the single languages. In fact this theory has been correctly rejected by renowned linguists, such as Schmitt (2000: 394) and Schlerath (1987, 1996). Furthermore, there is no archaeological substantiation for it either. The long led discussion about the Idg. social structure (which has been dealing merely with speculation) through constant repetition has given the impression that well-founded results do actually exist (Untermann 1989: 30). This being the situation, Zimmer (2004: 3) claims that: "Auf den Ansatz einer regelrechten, weitgehend systematischen 'Idéologie tripartie', wie er sie propagiert hat, wohl endgültig verzichtet werden muß". What has been stated thus far – that is: there is no sufficient evidence supporting the reconstruction of the social structure of the (assumed) Idg. community, in an assumed *Urheimat* – also applies to the religious conceptions, to the *Urreligion* of the Idg. This can be shown through an example drawn from the field of archaeology. The grave and burial customs of past times have always been seen as an expression of beliefs about the hereafter, the religious patterns of thought. Gimbutas (1980 and 1986) advocated an invasion of mounted eastern rider

nomads (Fig. 5) from the steppes of the north Pontic area up to the settlement area of the later Greeks, Celts, Germans, Balts, and Slavs (see Häusler 1998 and 2004), and constructed a strict contrast between the burial customs of the two reputedly clashing cultural worlds: on the one hand collective tombs with nameless dead, on the other hand prominent individual burials for the foreign conquerors. However, this is not at all correct (Häusler 2003 and 2004a: 201). In the Funnel Beaker Culture of the 4th millennium B.C., in the northern part of central Europe and in north Europe (Midgley 1992; Ebbesen 1998; Figs. 7, 8) existed on the one hand single graves equipped with specific grave goods, and on the other hand collective burials in megalithic tombs.

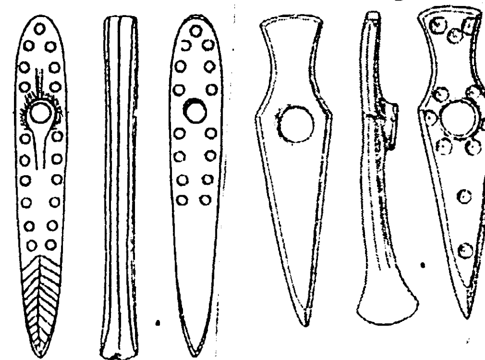


Fig. 7. Battle axes of the Funnel Beaker Culture in Scandinavia.

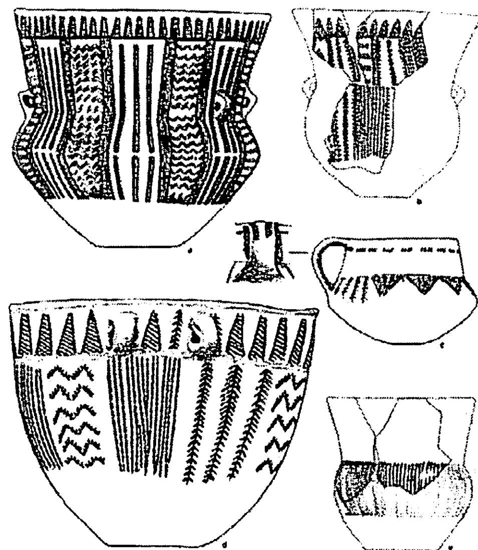


Fig. 8. Pottery of the Funnel Beaker Culture from Germany.

These possibly contained secondary burials of a socially elevated class of particularly esteemed ancestors. However, single burials provided with specific grave goods are not only typical of the Corded Ware Culture (Fig. 9), allegedly immigrating from the north Pontic area to central Europe, but were already widespread in Mesolithic Europe as well as in the Linear Pottery Culture and its successive cultures.

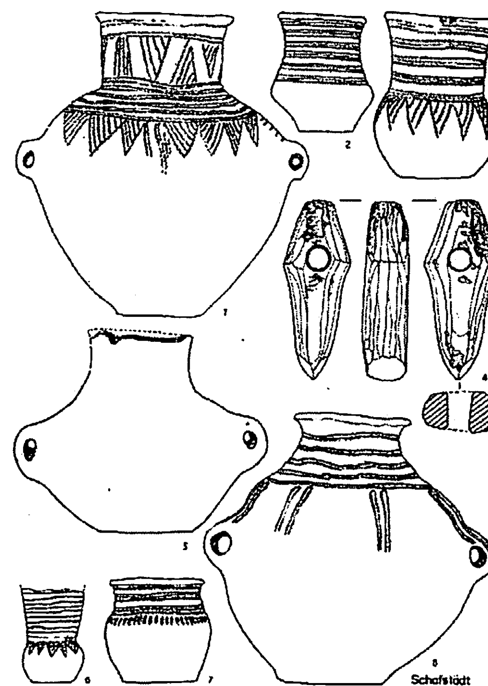


Fig. 9. Findings of central German Corded Ware.

Furthermore, there also existed serious differences. Corded Ware people of central Europe and the population of the Single Grave Culture buried their dead according to the principle of the 'bipolar gender differentiated burial' customs. In the central German Corded Ware Culture men are oriented to the west, crouched on their right side, whilst women are oriented to the east, crouched on their left side, both looking to the south (Fig. 10).

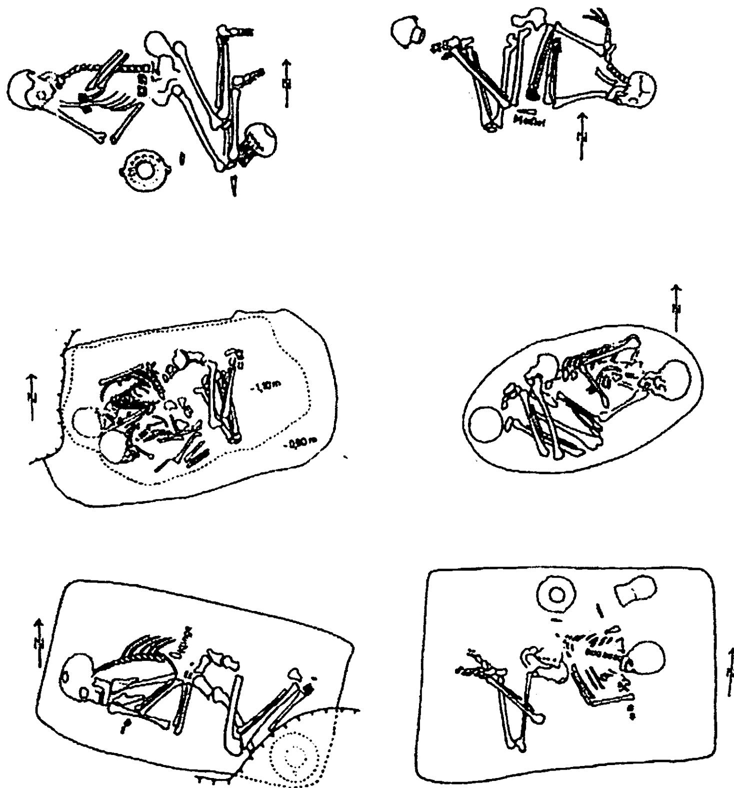


Fig. 10. Gender differentiated burial practices in the central German Corded Ware Culture: Males oriented to the west, crouched on their right side; females oriented to the east, crouched on their left side, both facing south.

However, this specific type of burial custom [which surely reflects equally specific beliefs about the status of man and woman in society and about the hereafter (Häusler 2003b and 2006)], is not attested in those eastern regions that are claimed to be the source of these campaigns of conquest, that is: in the older Ochre Grave Culture of the north Pontic area, as well as in Siberia and Central

Asia (at the time). Indeed, here we mainly find structures of burial customs in which such polarities were unknown. As an example of this, one can mention a few inhumation burials from a western offshoot of the Ochre Grave Culture in east Hungary (Fig. 4), which show west oriented corpses lying crouched on their backs, facing east. With respect to orientation rules, this burial custom can be regarded as 'monopolar', and with reference to the crouched position of the dead as 'gender in-differentiated'.

6. 2. The grave and burial customs in the region extending between the North Sea and Central Asia, in the post-mesolithic period, were regionally differentiated over millennia and subject to numerous, often fundamental structural innovations (Häusler 2004a). They cannot be traced back to a common denominator, a specific cultural structure which could be ascribed to a regionally restricted Idg. proto-culture, with specific burial customs as an expression of an equally specific Idg. proto-religion, within the area of an original homeland. Several scholars agree with this point of view. For example, Zimmer (2004:2) claims that, despite 200 years of research, it is still impossible to give a fairly coherent sketch of the "religion of the Uridg.". This is primarily due to the fragmentary character of all linguistic reconstructions, which do not go beyond an individual word or a short formula. Furthermore, modern accounts of this topic make assumptions for which there is no sufficient supporting evidence in the available, old texts of the Idg. languages. Schlerath (1998: 88) arrives at a similarly negative conclusion. This also applies to the assumption of a primary and a secondary homeland of the Idg., that is, of the original settlement areas of the Idg. in a 'final distribution area', at the time preceding the splitting off of the various Idg. branches (Häusler 2004: 84 ff.; Fig. 28).

I would like to summarize these results by simply stating that neither common beliefs regarding the hereafter, nor common

burial customs, nor a specific religion of the Idg., located in a primary homeland,or final distribution area, existed. This means that the thesis of an Idg. proto-culture, with a tangible social structure, religion and mythology, all confined in an original homeland, should be renounced.

7. The migration and invasion hypotheses

7. 1. The widespread distribution of the Idg. peoples and their language over large parts of Eurasia has been recognized, since early on, as a phenomenon that requires an explanation. The hypothesis of a catastrophe-like incursion, a full scale invasion of Europe, found special consensus (Fig. 5) and has proved to be particularly persistent. To justify the invasion / migration model all sorts of hypotheses were put forward. For instance, as triggering factors for the assumed migrations, some believed in the particular geographical conditions of the area, others in changes of climate or, often, in unusual, racial and psychical features and the superiority of the Idg. Already early on an effort was made to look for an original homeland of the Idg. in regions which were as unknown as possible, inhospitable or no longer accessible. Apart from Asia in general - Central Asia, the Pamirs, Hindu-Kush, Bactria, India, the Caucasus, Armenia, the Kyrgyz steppes as well as southern Russia, even the North Pole was declared to be the original homeland of the Idg. and their proto-language. Amongst others, Scherer (1968), de Benoist (1997), Mallory (1989), Beran (1998) and Raulwing (2000: 69ff.) inform us about the numerous (migrationist-type of) models of interpretation put forward to solve the Idg. problem. Within these models, hardly any region of northern Eurasia has been omitted, and even today, just as a century ago, there is no lack of attempts to declare the Funnel Beaker Culture of northern central Europe (Figs. 7, 8) as the starting-point of the Idg. migrations toward the east or south of

Europe [Boettcher (1999); see Häusler (2003a: 505)]. In order to understand the still currently popular eastern origin of the Idg., it is important to turn to the origin of this idea, which is to be found in the old biblical legend of the '*ex oriente lux*'. Herder thought of an origin of mankind in Asia, and also for Leibniz, Asia was the '*vagina populorum*' – these thoughts were quite natural in their time. Since Friedrich Schlegel, who in the year 1808 published his much-noticed text *Über die Sprache und Weisheit der Inder* in Heidelberg, there followed a fascination for Sanskrit which continues to have an effect to this day (Mayrhofer 1983 and 2004). For Schlegel it was clear that Sanskrit is older than Greek and Latin, never mind German and Persian. Sanskrit since then was interpreted as the oldest Idg. language stemming from India, and therefore as the ancient source and origin of all other Idg. daughter languages. Such opinions were held up to recent times. On the other hand, in archaeology, for a long time it has been almost natural to explain every conceivable appearance of a given culture through migrations from afar, and to think in terms of winners and losers, upper and lower strata, masters and slaves. Among others, Wahle (1932) and Güntert (1934) – both working in Heidelberg – have used these categories. According to them, hordes of military nomads from the north Pontic region would have embarked in migratory movements in order to subjugate the native population (Funnel Beaker Culture) and to impose the Corded Ware culture and the Idg. language upon them (see Häusler 1998: 2).

Gimbutas (1980 and 1986), who also spent some time in Heidelberg, propagated the same model in the Anglo-Saxon countries, although in a modified variant: she introduced the label 'Kurgan Culture'. According to this model, three waves of surging nomadic horsemen from the north Pontic steppes would have invaded and defeated the original, peaceful population of Old Europe, whereupon the Idg. language of the winners triumphed. The renowned historian Milojčić (1979: 243) described this

construct as: "An unproven research ghost with a long tradition". Schmoeckel (1985 and 2002) converted the ideas of Gimbutas into a novel-like presentation which in some circles of Indo-Germanists is still celebrated today as a revelation (according to Zeilfelder 2004). The Wahle-Güntert-Gimbutas concept is today used by some linguists in order to bestow a (pseudo-) archaeological support to their particular linguistic interpretation of Idg. Thus, for example, Beekes (1996) believes that there must have been an Old European – therefore not Idg. – language layer in Europe. In his opinion, the Srednij-Stog Culture of the 5th - 4th millennium B.C. (Fig. 11), which occupied a small region east of the lower Dnieper, would correspond best to the image of the early Idg. people (Beekes 1995: 50; see Häusler 2002, 2003: 53).

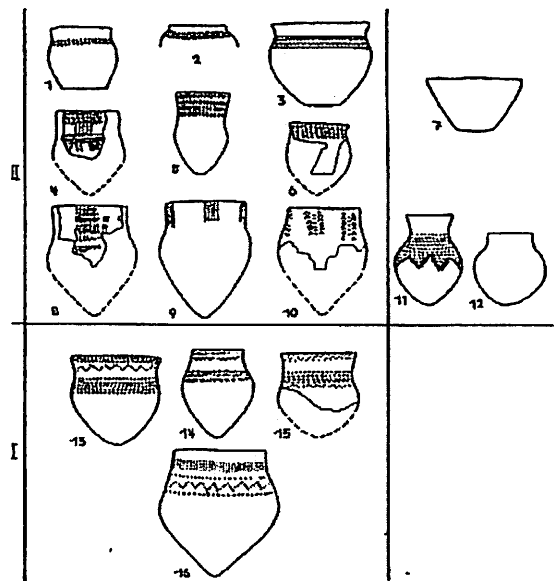


Fig. 11. Pottery of the Srednij-Stog Culture of the 5th – 4th millennium B.C., from Ukraine.

At their main site, Dereivka, some huts have been found. The population hunted wild horses but knew the domestic pig – not really good prerequisites for nomadic invaders. In the Corded Ware or Single Grave Culture (Fig. 9) Beekes sees the first Idg. people, who would have then advanced as far as the Netherlands. However, these views cannot be accepted, because none of the three Kurgan migratory waves, as assumed by Gimbutas, has been satisfactorily proved (see Häusler 1985). Furthermore, Govedarica (2004; see also Häusler 2006a) has demonstrated that the early Ochre graves of the north Pontic steppes (4650 - 4000 cal B.C.) cannot be explained through eastern invasions at all, as Gimbutas claimed, but have arisen on an autochthonous basis, under the influence of southeast European cultures of the eneolithic phase (particularly the Varna Culture of Bulgaria). Similarly, the Corded Ware or Single Grave Culture of central and northern Europe represents the result of continuous, progressive development of the local Funnel Beaker Culture to a Corded Ware lifestyle (Ebbesen 1997; Hübner 2004). Thus, the hypothesis of a pre-Idg. European substrate layer floats in a void. Further attempts to support controversial linguistic hypotheses by making reference to alleged waves of conquest from the north Pontic area have been discussed at length by Häusler (2003a, 2004: 19, 94). For instance, Fortson (2005: 40 ff.) uses Gimbutas' model as a mainstay of his own claim that the north Pontic area can be indeed identified with the homeland of the Idg.

Finally, I shall discuss the model of interpretation of the origin and spread of the Idg. as proposed by Gamkrelidze and Ivanov (1995). According to the Authors, the Idg. homeland would have lain at the upper reaches of the Euphrates and Tigris in the Halaf Culture (5th - 4th millennium B.C.). From that area, the Idg. would have first emigrated into Central Asia, to the Lake Aral, then they would have moved through the north Pontic area to finally arrive – as Germans, Celts and Balts – to the central and

northern parts of Europe (see Häusler 1998: 15). This hypothesis is largely based on Gimbutas' claims that there must have been waves of invasions from the north Pontic area during the Eneolithic time. However, as discussed above, Gimbutas' model cannot be proven archaeologically, and, as a consequence, Gamkrelidze and Ivanov's model too is doomed to failure. At this point it is important to highlight that all of the above mentioned migration and invasion hypotheses, whatever the direction of the migratory movements may be, cannot be confirmed in the light of archaeological and anthropological results. It is superfluous to examine further areas of northern Eurasia or Transcaucasia for possible departure-points of the assumed expansions. Ultimately – at least for the region between the North Sea and Caspian Sea – there is only evidence for a continual, progressive development of the old-established population and culture, undisturbed by migration, since the Mesolithic / Neolithic times up to the times of the ancient Greeks, Celts, Germans and Balts, as we know them from written sources.

7. 2. For some readers the suspicion could arise that archaeology is not capable of demonstrating the existence of contacts and cultural relations between different regions or cultures. However, this is not the case, as can be observed through the example of the two regions of central Europe and north Pontic. Here, there is the Globular Amphora Culture (4th / early 3rd millennium B.C.; see Beier (1988) and Fig. 12), to which all along cultural influences from long distance migrations between east and central Europe have been attributed (see Häusler 1994, 2004: 66).

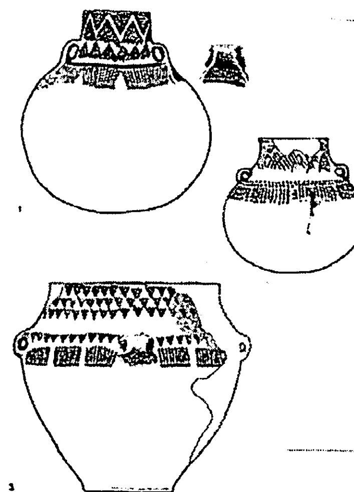


Fig. 12. Findings of the Globular Amphora Culture from Sargstedt, central Germany.

The Majkop Culture (3950 - 3300 cal B.C.), which emerged in the north Caucasus mainly on an autochthonous basis, and the Globular Amphora Culture, extending between Germany and Ukraine (more precisely, between Lower Saxony in the northwest as well as Wolhynian and Podolian in the southeast), are not derivable from each other, neither were there interrelations between them. Szmyt (1996) has looked more closely into the contacts between the Globular Amphora Culture, autochthonous in central Europe, and the Ochre Grave Culture of the north Pontic area. There have been contacts between these two cultures in a peripheral area, and these are reflected in decorative patterns, imported flint adzes in the Ochre Grave Culture; occasionally also influences in the way tombs are constructed may be observed. In this case, cultural influences ran mainly from west to east and not from east to west. These marginal cultural relations show that we are dealing in general with two isolated culture blocks in Eastern

and central Europe. The same applies also to the period of the Corded Ware Beaker cultures of Europe, in the 3rd millennium B.C. In contrast to these marginal cultural relations between east and central Europe, the influences which spread from southeast Europe or the Balkans into central and northern Europe over millennia, since the Neolithic, had a significantly stronger dimension.

This being the case, the thesis of the repeatedly evoked, cross-border cultural flow, of the far-reaching migration movements that brought into motion new groups of people together with their new, Idg. language, can be eliminated from the start.

8. The model of Lord C. Renfrew

In this paragraph I shall present the model proposed by Lord C. Renfrew (1987; modified version 1999; final version 2003 and 2003a). Renfrew assumes that languages often spread or are adopted in connection with new cultural achievements, if they are perceived to be of higher prestige. He links the spread of the Idg. languages with the dispersal of neolithic people and the neolithic way of life in Europe, starting out from Anatolia in the 7th millennium B.C., a way of life which was far superior to that of the local hunters, gatherers and fishermen (Renfrew 2003; Fig. 2-6). Even if the neolithisation of Europe was not carried out by an extensive immigration of sections of the population originating from the south, but only by a low number of settlers at a time, this would have sufficed, according to Renfrew, to establish a new economic system and a new language system in Europe. The spread of this new language cannot have been a linear process; on the contrary, it would have developed over a long period of time, such as in the area of the early Neolithic cultures of the Balkans or, more generally, of south east Europe. From the eastern foothills of the aforementioned area (Tripolye Culture of Ukraine) the Idg.

language would have been finally adopted also in the north Pontic steppes (where Scythians and Sarmathians were found, who spoke an Indo-Iranian language). This model of interpretation, which presupposes the existence of a (controversial) Indo-Hittite phase, was largely rejected by Indo-Germanists [see the register of the reviews of Renfrew's monograph from the year 1987; see also Raulwing (2000: 133)]. In fact, if one accepts the mainstream tenets of Idg., it is hard to imagine the formation of an Indo-Hittite or Idg. linguistic community already in the 7th millennium B.C. In addition, this model presupposes that the Hittites, attested at the end of the 2nd millennium B.C., were direct descendants of the Indo-Hittite community in Anatolia, whilst the main stream doctrine is that the Hittites emigrated from a still unknown region from the North only in the 2nd millennium B.C. The terminology for the 'Wheel Axle Complex' is regarded as an integral component of the Idg. parent language, reconstructed with the help of linguistic palaeontology. Apart from the question whether wagons with disc wheels drawn by oxen were invented in the Near East, in Europe or, 'polycentrically', independently from one another, this innovation is first documented from the middle to end of the 4th millennium B.C. (Fansa & Burmeister 2004; Schönfelder 2005; Häusler 2004). Idg. speaking population groups of the 7th millennium B.C., of course, could not have spread together with the lexicon referring to an invention which was made only millennia later. Here Renfrew argues that Idg. is characterised primarily through its linguistic structure unique in Eurasia, whilst the terms relating to cultural achievements could be of secondary nature. In other words, we would be facing here a typical issue of 'depth of time' (Renfrew, McMahon & Trask 2000), that is, some layers of Idg. vocabulary may well be younger than others.

A similar interpretation of the 'Idg. problem' is advocated by K. Wiik (2000 and 2002) with regard to the Finno-Ugric peoples. According to the Author, large parts of Europe were

formerly occupied by bearers of Finno-Ugric languages. Over the course of the millennia these populations adopted the Idg. language, the new *lingua franca* – also perceived as being of higher prestige – together with the neolithic way of life of the newcomers. Linguistic differences (i.e., different varieties of the Idg. languages) would have arisen because of the imperfect learning of the foreign idiom by the old-established, pre-Idg. populations.

9. Conclusion

In the above contribution only a few aspects could be pointed out which are discussed in connection with the problem of the origin and spread of the Idg. languages and their bearers. It was mainly about drawing attention to some alternative explanatory models. The claim of the close affiliation among the Idg. languages is unquestionable. The claim that their roots reach far back into the past is equally beyond doubt. However, the problem of their origin cannot be solved by the widespread hypotheses about conquerors coming from the East. In fact, there is no evidence for any migrations from the East within the Neolithic or Copper Age. On the contrary, at least for the region extending between the North Sea and Caspian Sea, there is only evidence for a continual, progressive development of the old-established population and culture, undisturbed by migration, since the Mesolithic / Neolithic times up to the times of the ancient Greeks, Celts, Germans and Balts.

As already mentioned in the abstract, the diffusionist theses do not appear to contradict the archaeological evidence; still, the Idg. linguistic community is not identifiable by means of archaeological or anthropological research.

References

Beekes, R.S.P.
1995 *Comparative Indo-European Linguistics. An Introduction.* Amsterdam / Philadelphia: Benjamin.
1996 Ancient European Loanwords. *Historische Sprachforschung* 109: 215-236.

Beier, H. J.
1988 Die Kugelamphorenkultur im Mittelelbe-Saale-Gebiet und in der Altmark. *Veröffentlichungen des Landesmuseums für Vorgeschichte Halle* 41. Berlin: VEB Deutscher Verlag der Wissenschaften.

de Benoist, A.
1997 Indo-européens: à la recherche du foyer d'origine. *Nouvelle Ecole* 49:13-105.

Beran, J.
1998 Zur Sprache im mitteleuropäischen Neolithikum. In: Preuß, J. (ed.), *Das Neolithikum in Mitteleuropa. Kulturen - Wirtschaft - Umwelt vom 6. bis 3. Jahrtausend v. u. Z.*, Bd. 1/1. Teil A: *Das Neolithikum in Mitteleuropa.* Weissbach: Beier & Beran. 242-256.

Boettcher, C.-H.
1999 *Der Ursprung Europas. Die Wiege des Westens vor 6000 Jahren.* St. Ingbert: Röhrig Universitätsverlag.

Bopp, F.
1816 *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persichen und germanischen Sprache.* Frankfurt am Main: Andreä.

Dixon, R. M.W.
1997 *The Rise and Fall of Languages.* Cambridge: Cambridge University Press.

Dumézil, G.
1958 *L'idéologie tripartie des Indo-Européens.* Paris: Berchem Latomus.

Ebbesen, K.
1997 Der Beginn der Streitaxtzeit. In: *Early Corded Ware Culture. The A-Horizon – fiction or fact? International Symposium in Jutland 2nd - 7th May 1994. Arkæologiske Rapporter* 2: 75-91. Esbjerg: Esbjerg Museum.

Ebbesen, K.
1998 Frühneolithische Streitäxte. *Acta Archaeologica* (København) 69: 77-112.
E hret, Ch.
1975 Linguistic evidence and its correlation with archaeology. *World Archeology* 8. No 1: 5-18.
Fansa, M. and S. Burmeister (eds)
2004 *Rad und Wagen. Der Ursprung einer Innovation. Wagen im Vorderen Orient und Europa.* Beiheft der *Archäologische Mitteilungen aus Nordwestdeutschland* 40. Mainz: Philipp von Zabern.
Feist, S.
1926 Indogermanen. In: M. Ebert (ed.), *Reallexikon der Vorgeschichte* Bd. 6: 54-66. Berlin
Fortson IV, B.W.
2005 *Indo-European Language and Culture. An Introduction.* Oxford: Blackwell Publishing.
Gamkrelidze, Th.V. and V.V.Ivanov
1995 *Indo-European and the Indo-Europeans, Part I.* Berlin and New York: Mouton de Gruyter.
Gimbutas, M.
1980 The Kurgan Wave No. 2 (ca. 3400 - 3200 B.C.) into Europe and the Following Transformation of Culture. *Journal of Indo-European Studies* 8: 273-317.
1986 Remarks on the Ethnogenesis of the Indo-Europeans in Europe. In: Bernhard, W. and A. Kandler-Pálsson (eds), *Ethnogenese europäischer Völker*. Stuttgart and New York: Gustav Fischer. 5-20.
Govedarica, B.
2004 *Zepterträger - Herrscher der Steppen. Die frühen Ockergräber des älteren Äneolithikums im karpatenbalkanischen Gebiet und im Steppenraum Südost- und Osteuropas.* Heidelberger Akademie der Wissenschaften. Internationale Interakademische Kommission für die Erforschung der Vorgeschichte des Balkans. Monographien Bd. 6. Harald Hauptmann (ed.). Mainz: Philipp von Zabern.
Güntert, H.
1934 *Der Ursprüng der Germanen.* Heidelberg: Carl Winter

Häusler, A.
1985 Kulturbeziehungen zwischen Ost- und Mitteleuropa im Neolithikum? *Jahresschrift für Mitteldeutsche Vorgeschichte* 68: 21-74.
1994 Die Majkop-Kultur und Mitteleuropa. *Zeitschrift für Archäologie* 28: 191-246.
1998 Zum Ursprung der Indogermanen. Archäologische, anthropologische und sprachwissenschaftliche Gesichtspunkte. *Ethnographisch-Archäologische Zeitschrift* 39: 1-46.
2000 Review of J.P. Mallory and D.Q. Adams, *Encyclopedia of Indo-European Culture*, London / Chicago: Fitzroy Dearborn, 1997. *Indogermanische Forschungen* 105: 314-318.
2002 Ursprung und Ausbreitung der Indogermanen: Alternative Erklärungsmodelle. *Indogermanische Forschungen* 107: 47-75.
2003 Urkultur der Indogermanen und Bestattungsriten. In: A. Bammesberger and Th. Vennemann (eds), *Languages in Prehistoric Europe.* Heidelberg: Universitätsverlag Winter. 49-83.
2003a Über alte und neue Hypothesen zum Ursprung und zur Verbreitung der Indogermanen. *Ethnographisch-Archäologische Zeitschrift* 44: 503-521.
2003b Geschlechtsdifferenzierte Bestattungssitten und die Entstehung des grammatischen Geschlechts in den indogermanischen Sprachen. In: E. Jerem and P. Raczky (eds), *Morgenrot der Kulturen. Frühe Etappen der Menschheitsgeschichte in Mittel- und Südosteuropa. Festschrift für Nándor Kalicz zum 75. Geburtstag. Archaeolingua* 15: 39-52. Budapest: Archaeolingua Alapitvány.
2004 Nomaden, Indogermanen, Invasionen. Zur Entstehung eines Mythos. *Orientwissenschaftliche Hefte* 5/2003. Mitteilungen des SFB "Differenz und Integration" 3, 2nd edition. Halle: Orientwissenschaftliches Zentrum der Martin-Luther-Universität Halle-Wittenberg
2004a Überlegungen über Systematisierung, Analyse und Interpretation von Körperbestattungen des Neolithikums, Neolithikums und der Bronzezeit. *Jahresschrift für Mitteldeutsche Vorgeschichte* 88: 149-217.

2006 Polaritäten, Geschlechtsdifferenzierte Bestattungssitten und der Ursprung des grammatischen Geschlechts in den indogermanischen Sprachen. *Jahreschrift für Mitteldeutsche Vorgeschichte* 90: 71-150.

2006a Review of B. Govedarica, *Zepterträger – Herrscher der Steppen. Die frühen Ockergräber des älteren Äneolithikums im karpatenbalkanischen Gebiet und im Steppenraum Südost- und Osteuropas*. Heidelberger Akademie der Wissenschaften. Internationale Interakademische Kommission für die Erforschung der Vorgeschichte des Balkans. Monographien Bd. *6*. Harald Hauptmann (ed.). Mainz: Philipp von Zabern. *Indogermanische Forschungen* 11: 346-349.

Hübner, E.
2004 Jungneolithische Gräber auf der Jütischen Halbinsel. Typologische und chronologische Studien zur Einzelgrabkultur. *Archäologisches Nachrichtenblatt* 9: 214-219.

Khazanov, A.M.
1994 *Nomads and the Outside World*. Wisconsin: The University of Wisconsin Press.

Klaproth, J.
1823 *Asia polyglotta*. Paris: A. Schubart.

Koerner, E.F.K.
1982 Observations on the Sources, Transmission and Meaning of 'Indo-European' and Related Terms in the Development of Linguistics. *Indogermanische Forschungen* 86: 1-29.

Koivulehto, J.
1995 Finnland. In: *Reallexikon der Germanischen Altertumskunde*, 2nd edition, Bd. 9: 77-89. Berlin / New York.

2003 Frühe Kontakte zwischen Uralisch und Indogermanisch im nordwestindogermanischen Raum. In: A. Bammesberger and Th. Vennemann (eds), *Languages in Prehistoric Europe*. Heidelberg: Universitätsverlag Winter. 279-317.

Lüning, J.
2000 *Steinzeitliche Bauern in Deutschland. Die Landwirtschaft im Neolithikum*. Bonn: Dr. Rudolf Habelt GmbH .

Mallory, J.P.
1989 *In Search of the Indo-Europeans*. London: Thames and Hudson.

Mallory, J.P. and D.Q. Adams (eds)
1997 *Encyclopedia of Indo-European Culture*. London / Chicago: Fitzroy Dearborn.

Matthias, W
1982 *Kataloge zur Mitteldeutschen Schnurkeramik. T. V. Mittleres Saalegebiet.* Veröffentlichungen des Landesmuseums für Vorgeschichte Halle 35. Berlin: VEB Deutscher Verlag der Wissenschaften.

Marcantonio, A.
2002 *The Uralic Language Family: Facts, Myths and Statistics*. Oxford / Boston: Blackwell.

Mayrhofer, M.
1983 Sanskrit und die Sprachen Alteuropas. Zwei Jahrhunderte des Widerspiels von Entdeckungen und Irrtümern. *Nachrichten der Akademie der Wissenschaften Göttingen* 5:121-153.

2004 Über Friedrich Schlegel und die Anfänge der Indogermanistik. *Die Sprache* 43, H. 2, 2002/03: 221-232.

Meier-Brügger, M.
2002 *Indogermanische Sprachwissenschaft*. 8th edition. Berlin / New York: Walter de Gruyter.

Midgley, M. S.
1992 *TRB culture. The First Farmers of the North European Plain.* Edinburgh: Edinburgh University Press.

Milojčić, V.
1979 Review of A. Tovar, Krahes, *Alteuropäische Hydronymie und die westindogermanischen Sprachen*, Heidelberg 1977. *Prähistorische Zeitschrift* 54: 243.

Mühlmann, W. E.
1968 *Geschichte der Anthropologie*. Bonn: Atheneüm.

Nuñez, M.
1998 Old and new ideas about the origin of the Finns and Saami. In: K. Julku and K.Wiik (eds), *The Roots of Peoples and Languages of Northern Eurasia I*. Turku: Societas Historiae Fenno-Ugricae. 151-167.

2000 Problems with the Search for the Ancestral Finns. In: A. Künnap (ed.), *The Roots of Peoples and Languages of Northern Eurasia II and III*. Tartu: University of Tartu. 60-68.

2002 The rise and fall of Hackman's archaeological-linguistic model and its aftermath: An epistemological appraisal. In: K. Julku

(ed.), *The Roots of Peoples and Languages of Northern Eurasia IV.* Oulu: Societas Historiae Fenno-Ugricae. 197-212.

Probst, E.
1991 *Deutschland in der Steinzeit.* München: C. Bertelsmann.

Pulgram, E
1968 Indoeuropäisch und "Indoeuropäer". In: A. Scherer (ed.), *Die Urheimat der Indogermanen. Wege der Forschung 116.* Darmstadt: Wissenschaftliche Buchgesellschaft. 455-477.

Raulwing, P.
2000 Horses, Chariots and Indo-Europeans. Foundations and Methods of Chariotry Research from the Viewpoint of Indo-European Linguistics. *Archaeolingua. Series Minor* 13. Budapest: Archaeolingua Alapitvány.

Renfrew, C.
1987 *Archaeology and Language: The Puzzle of Indo-European Origins.* London: Jonathan Cape.
1999 Time Depth, Convergence Theory, and Innovation in Proto-Indo-European: 'Old Europe' as a PIE Linguistic Area. *Journal of Indo-European Studies* 27, No. 3-4: 257-293.
2003 Convergence Theory, and Innovation in Proto-Indo-European: 'Old Europe' as a PIE Linguistic Area. In: A. Bammesberger and Th. Vennemann (eds), *Languages in Prehistoric Europe.* Heidelberg: Universitätsverlag Winter. 17-48.
2003a The Neolithic Transition in Europe: Linguistic Aspects. In: A. J. Ammerman and P. Biagi (eds), *The Widening Harvest. The Neolithic Transition in Europe: Looking Back, Looking Forward.* Boston: Archaeological Institute of America. 327-336.

Renfrew, C., McMahon, A. and L. Trask (eds)
2000 *Time Depth in Historical Linguistics.* Cambridge: McDonald Institute for Archaeological Research.

Robb, J.
1991 Random causes with directed effects: the Indo-European language spread and the stochastic loss of lineages. *Antiquity* 65: 287-291.
1993 A social prehistory of European languages. *Antiquity* 67: 747-760.

Scherer, A. (ed.)
1968 Die Urheimat der Indogermanen. *Wege der Forschung* 116. Darmstadt: Wissenschaftliche Buchgesellschaft.

Schlegel, F.
1808 *Über die Sprache und Weisheit der Inder. Ein Beitrag zur Begründung der Alterthumskunde.* Heidelberg: Mohr & Zimmer.

Schleicher, A.
1861 *Compendium der vergleichenden Grammatik der indogermanischen Sprachen.* Weimar: H. Böhlau.

Schlerath, B.
1973 Die Indogermanen. Das Problem der Expansion eines Volkes im Lichte seiner sozialen Struktur. *Innsbrucker Beiträge zur Sprachwissenschaft. Vorträge* 8. Innsbruck: Institut für Vergleichende Sprachwissenschaft der Universität Innsbruck.
1981 Ist ein Raum/Zeit-Modell für eine rekonstruierte Sprache möglich? *Zeitschrift für vergleichende Sprachwissenschaft* 95: 175-202.
1987 Können wir die urindogermanische Sozialstruktur rekonstruieren? Methodologische Erwägungen. In: W. Meid (ed.), *Studien zum indogermanischen Wortschatz. Innsbrucker Beiträge zur Sprachwissenschaft* 52: 249-264. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck
1992 Review of W. Meid, *Archäologie und Sprachwissenschaft. Kritisches zu neueren Hypothesen der Ausbreitung der Indogermanen.* Innsbruck 1989. Institut für Sprachwissenschaft der Universität Innsbruck. *Prähistorische Zeitschrift* 67: 137-139.
1996 Georges Dumézil und die Rekonstruktion der indogermanischen Kultur. 2. Teil. *Kratylos* 41:1-67.
1998 Religion der Indogermanen. In: W. Meid (ed.), *Sprache und Kultur der Indogermanen.* Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck. 87-99.

Schmid, W.P.
1968 *Alteuropäisch und Indogermanisch.* Wiesbaden: Franz Steiner Verlag.
1983 Das sprachgeschichtliche Problem Alteuropa. *Sprachwissenschaft* 8: 101-113.
1983a Griechenland und Alteuropa im Blickfeld des Sprachhistorikers. *Epistemonike epeterida tes filosofikes scholes tou Aristoteleion Panepistemoniou Thessalonikes* 21: 397- 412.

1985 Das Lateinische und die Alteuropa-Theorie. *Indogermanische Forschungen* 90:129-146.
1987 'Indo-European' - 'Old European' (On the Reexamination of Two Linguistic Terms). In: S.N. Skomal and Ed. C. Polomé (eds), *Proto-Indo-European: The Archaeology of a Linguistic Problem*. Washington: Institute for the Study of Man. 322-338.
1995 Alteuropäische Gewässernamen. In: E. Eichler and G. Hilty (eds), *Namenforschung. 1. Teilband*. Berlin / New York: Walter de Gruyter. 756-762.
1998 Methodische Bemerkungen zur Klassifikation: Alteuropäisch. *Onomastica Slavogermanica* 23: 21-28.

Schmitt, R.
2000 Indogermanische Altertumskunde. I Sprachliches. In: *Reallexikon der Germanischen Altertumskunde*, 2nd edition, Bd. 15: 384-402. Berlin / New York.

Schmitt, R. and A. Häusler
2000 Indogermanische Altertumskunde. In: *Reallexikon der Germanischen Altertumskunde*, 2nd edition, Bd. 15: 384-408. Berlin and New York.

Schmoeckel, R.
1985 *Die Hirten, die die Welt veränderten*. Hamburg: Rheinbeck.
2002 *Die Indoeuropäer. Aufbruch aus der Vorgeschichte*. 2nd edition. Bergisch Gladbach: Verlagsgruppe Lübbe.

Schönfelder, M.
2005 Review of Fansa, M. and S. Burmeister (eds), *Rad und Wagen. Der Ursprung einer Innovation. Wagen im Vorderen Orient und Europa*. Beiheft der *Archäologische Mitteilungen aus Nordwestdeutschland* 40. Mainz: Philipp von Zabern. *Germania* 83: 503-507.

Seebold, E.
1998 Germanen B. Sprache und Schrift. In: *Reallexikon der Germanischen Altertumskunde*, 2nd edition, Bd. 11: 275-305. Berlin / New York.

Shapiro, F.R.
1981 On the Origin of the Term 'Indo-Germanic'. *Historiographia Linguistica* 8:165-170.

Shilov, V.P.
1989 The origins of migration and animal husbandry in the steppes of eastern Europe. In: J. Clutton-Brock (ed.), *The walking larder. Patterns of Domestication, Pastoralism, and Predation*. Boston / Sidney: University Press. 119-126.

Sutrop, U.
2000 The Forest of Finno-Ugric Languages. In: A. Künnap (ed.), *The Roots of Peoples and Languages of Northern Eurasia II and III*. Tartu: University of Tartu. 165-196.

Szmyt, M.
1996 Globular Amphora Culture in Eastern Europe. Present State of Research and Possibilities for Future Studies. *Baltic-Pontic Studies* 4:3-27.

Trubetzkoy, N.S.
1968 Gedanken über das Indogermanenproblem. In: A. Scherer (ed.), *Die Urheimat der Indogermanen. Wege der Forschung 116*. Darmstadt: Wissenschaftliche Buchgesellschaft. 214-223.

Udolph, J.
1994 *Namenkundliche Studien zum Germanenproblem*. Berlin / New York: Carl Winter Universitätsverlag.

Untermann, J.
1985 Ursprache und historische Realität. Der Beitrag der Indogermanistik zu Fragen der Ethnogenese. In: *Studien zur Ethnogenese* 7: 133-163. Opladen: Rheinisch-Westfälische Akademie der Wissenschaften.
1989 Review of W. Meid (ed.), Studien zum indogermanischen Wortschatz. *Innsbrucker Beiträge zur Sprachwissenschaft* 52. Innsbruck 1987. *Kratylos* 34: 45-54.

Wahle, E.
1932 *Deutsche Vorzeit*. Basel: Schwabe.

Wiik, K.
2000 European lingua francas. In: A. Künnap (ed.), *The Roots of Peoples and Languages of Northern Eurasia II and III*. Tartu: University of Tartu. 202-236.
2002 On the Emergence of the Main Indo-European Language Groups of Europe through Adstratal Influence. In: K. Julku (ed.), *The Roots of Peoples and Languages of Northern Eurasia IV*. Oulu: Societas Historiae Fenno-Ugricae. 285-292.

Zeilfelder, S.
2004 Review of R. Schmoeckel, *Die Indoeuropäer. Aufbruch aus der Vorgeschichte*. 2nd edition. Bergisch Gladbach: Verlagsgruppe Lübbe. *Kratylos* 49: 177.

Zimmer, S.
2002 Tendenzen der Indogermanischen Altertumskunde 1965-2000. I. Teil: Sachkultur. *Kratylos* 47:1-22.
2004 Tendenzen der Indogermanischen Altertumskunde 1965-2000. II. Teil: Geistige Kultur. *Kratylos* 49: 1-25.

References for the plates and figures

1	after Seebold 1998, p. 283
2	after Sutrop 2000, Fig. 9
3a-d	after Robb 1991, Fig. 1-4
4	after Häusler 2004, Fig. 17
5	after Probst 1991, p. 401
6	after Raulwing 2000, Fig. 20
7	after Ebbesen 1998, Fig. 4
8	after Midgley 1992, Fig. 27
9	after Matthias 1982, Tab. 101
10	after Häusler 2003b, Fig. 4
11	after Häusler 2004, Fig. 6
12	after Beier 1988, Tab. 27

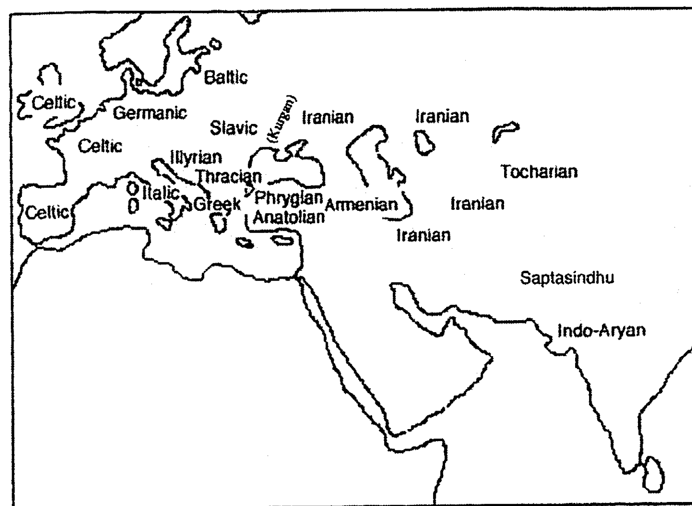
I thank the Authors and Editors for granting me copy-right permission of reproducing the listed tables and figures. (A.H.)

Indo-European Linguistics and Indo-Aryan Indigenism.

N. Kazanas
Omilos Meleton, Athens.

0. Abstract. This paper examines the general IE[1] issue and argues in favour of Indoaryan indigenism against the AIT (Aryan Invasion/Immigration Theory) which has been mainstream doctrine for more than a century. The extreme positions that there was no PIE[1] language or that this language is as currently reconstructed are refuted: the evidence suggests there was a PIE language but this cannot be reconstructed and all efforts in this reconstruction are misplaced. Since they are in no way verifiable, they should not be used as evidence for historical events. It is admitted even by rabid Indian nationalists that humans came to India from Africa sometime in the Pleistocene, and now there is evidence of change in the skeletal record of the region indicating that a new people may have entered c 6000-4500[2]; even so, if these people were the IAs, they must, surely, be regarded as indigenous by 1700. Recent genetic studies do not suggest any entry of IAs within the last 10 000 years but state that the European peoples came out of South Asia after 50 000 BP. Apart from such studies, other kinds of evidence and arguments will be used in full to demonstrate indigenism.

Indoeuropean and ProtoIndoEuropean



Map of IE languages

The map shows the location of languages broadly, not peoples or races; it is a modification of the map used by J. V. Day, 2001, p xiii.

1. Let us first see the ensuing discussion in its natural historical context.

American archaeologist J. Shaffer had the courage to call the AIT of India "a myth" (1984). The development of this "myth" which had obtained mainstream status in academia is well traced by E. Bryant (2001), J. Day (1994) and J. Mallory (1973). Here suffice it to say that having started as a linguistic theory, it soon acquired biological undertones involving more or less obvious ethnic/racial prejudices (Bryant 2001; Trautmann 1997; Shaffer 1984). Before the Nazi "aryanism" of the 1930's, the AIT was used by colonial politics as is obvious in British Prime Minister Baldwin's speech in Parliament in 1929: " Now after ages ... the two branches of the Aryan ancestry have again been brought together by Providence ... By establishing British rule in India,

God said to the British, 'I have brought you and the Indians together after a long separation ... it is your duty to raise them to your level as quickly as possible ... brothers that you are'"! God's ways were no longer so mysterious.

The linguistic theory which c 1800 CE sought to account for the similarities between Sanskrit, Greek, Latin, Germanic etc, contained the notion of an IE family of languages and a PIE mother-tongue. The literature on the subject is increasing enormously year by year. Hereafter I shall refute first the rejection by some scholars of a PIE tongue and even a family of IE languages and then the claims of others about the nature of PIE and its alleged "reconstruction". There is sufficient evidence to show that there had been a unitary PIE language and civilization but not enough to enable us to go very far with reconstructions. In any case, in the absence of people speaking PIE and of recognizably original texts, the reconstruction of the language – even if it were possible, which it is not – would be pointless and wasteful of other people's money. E. Leach wrote that after the discovery of the ISC (=Indus-Sarasvati Civilization) "Indo-European scholars should have scrapped all their historical reconstructions and started again from scratch. But this is not what happened. Vested interests and academic posts were involved" (1990).

'Vested' interests and academic posts are still involved and will continue to be involved because the human ego is not educated to let go of claims that are shown to be untrue, despite much trumpeting to the contrary. At the same time, although this phenomenon may cause and has caused considerable delay and darkness on the path of progress, it cannot ultimately prevent the establishment of true knowledge. A good case in point is the so-called "Copernican revolution" which, in general terms, ousted the Ptolemaic geocentric system of the heavens that had beclouded knowledge for some 18 centuries since Aristotle, and brought the heliocentric system which had been known by Greek naturalists (especially Aristarchos of Samos) even before Aristotle (Kuhn 1970, Koestler 1954). For many decades after the publication of Copernicus's work *De Revolutionibus* ... in 1543 the

mainstream savants of the day concocted more and more epicycles of the planets to account for the mathematics involved and the new observations. Eventually, after about a century the old model collapsed: more and more naturalists abandoned it in favour of the new heliocentric system. Similar concoctions, like the astronomical epicycles, are being produced in the field of Indology but with the new genetic studies (§19) the issue of indigenism can no longer be easily refuted.

2. The IndoEuropean family. That the IE family of languages is not a "family" is by no means a recent view, though it is not as old as the initial hypothesis that there is a "family" and a unitary PIE civilization and language. We love to speculate and set up theories and models – the more elegant the better. The human mind does not ordinarily seem to like gaps and emptiness. So it hastens to fill a gap in knowledge with some hypothesis. Then others raise objections and set up a different or even an opposite theory and so on. Thus, we speculate about everything raising one theory after another to keep mind and pen occupied. Here is an example out of myriads. A. Speirs writes (1978:26) "The principle of economy of hypothesis requires us to assume or at least to test the proposition that the labio-velars were PIE phonemes which in the PIE period shifted to labials and dentals in some circumstances, lost the labial element in others..." etc. We have the theory of a PIE, then the theory of the existence of labio-velar phonemes (§6, end), then speculation about what happened to these imaginary entities. Of course theories are useful because they become a tool for further research so long as one does not cling to the theory and argue that this is the one and only truth. Thus from the first half of the 20th c. several scholars began to doubt the existence of PIE. I give a selection of them: Bonfante 1931; Walder 1936; Trubezkoy 1939; Pisani 1949/71; Morgan Keeley 1992; Marcantonio 2002/05. There are many more, all doubting that what are described as IE branches (=Sanskrit, Greek, Germanic, etc) are offshoots of a unitary PIE. One of the explanations is that the similarities in these so-called IE branches are due to chance and have no further significance. Now, some fortuitous similarities in lexical items are not merely possible but

quite probable – and this could apply to languages outside the IE family. But we are hardly justified in evoking "chance" or "accident" when we meet 25 undoubtedly cognate stems present in three or more IE branches across Eurasia and designating the very same parts of the human body: e.g. eye, nose, tooth, jaw, breast, knee, foot, flesh, heart, blood etc. We find also at least 15 stems of equally certain cognation in three or more branches denoting relations and functions like brother, chief, carpenter/ fashioner, daughter, father, husband/master, man, mother etc (see §3, b). Then could it possibly be accidental similarity that we find across Eurasia but not in the N-East S *lobha* 'longing', L *lu-/li-bido* 'desire', Gmc *lufu/liubi* and Sl *ljubi* 'love' or S *mās* Av *māh-*, Toch A *mañ*, Gmc *mona*, B *menuo* Sl *meis-ic* all 'moon' and Gk *meis/mēn*, L *mens-is* and C *mī* 'month'?

Perhaps, yes, it may be, as others have suggested, that borrowing or waves of diffusion, sporadic or whatever, or convergence through long/repeated contact, are responsible for the similarities and cognations (Trubezkoy 1939). But how probable is this? ... It may be probable that 10, 20, 50 perhaps lexical items of everyday common use travelled across Eurasia at different times and in different directions. But this involves grave difficulties since the similarities and cognations in three or more IE branches run into many hundreds. Beyond this, there are close similarities in inflexion and conjugation and also in many social and cultural aspects (always exclusive of other non-IE peoples). Moreover, these similarities are not evident at all in Near Eastern languages.

Another objection raised is that common grammatical features appear only in S, Av, Ht, Gk and Italic. These branches, it is said, formed a closely related group that developed these common features. This is fallacious because, to take just one isogloss, while S and Av are satem, Ht, Gk and Italic are *kentum*. Moreover, as we shall see in §3 both lexical items and grammatical features (nominal and verbal endings, reduplication, etc) are spread right across the IE spectrum and are not confined to the group of S, Ht, Gk and Italic.

3. Let us look in detail at some of the difficulties involved in diffusion-waves and the close group.

(a) All Near-Eastern languages are - but for negligible exceptions which are easily accountable as borrowings - unaffected by these hypothetical waves or long/repeated contacts. Yet, roughly speaking, they are located in between Tocharian, Sanskrit and Iranian in the east and Greek, Italic, Germanic, Baltic and Celtic in the west. Some studies that claim elements "shared by Indo-European and Semitic" (eg Levin 1991) and Greek and Semitic (Burkert 1992) show very clearly that the "shared" material is utterly insignificant in comparison with the correspondences in IE branches.

(b) The obviously cognate words (=similar in sound and sense) in three or more IE branches are far too many for random or deliberate borrowing. Taking these very criteria, i.e. presence in three branches and designation of more or less the same (relatively) invariable entity, quality or activity, I easily collected about 500 stems (Kazanas 2007b). Some are common to all or most branches and many to more than three. E.g.:-

1. arm: S *bāhu*; Av *bāzu;* Gk *pēchus*; Gmc *buog*; Toch A *poke.*
2. belly: S *udara-*; Av *udara-*; Gk *hoderos* (=gaster); L *uterus* (venter?); B *vēderas.*
3. brother: S *bhrātṛ*, Av *brātār-*; Gk *phratēr* (= member of brotherhood); L *frāter*; C *brathir*; Gmc *bruodar*; B *broter-*; Sl *bratrŭ*; Toch A *pracar*.
4. clan, tribe: S *jana/jāti*; Av *-zana-/-zantiš* ; Gk *genos (phulē)*; L *gens (tribus)*; Gmc *cynn/kyn.*
5. to desire/love: S *lubhyati*; L *lub-/lib-et*; Gmc *liufs/lufaþ*; Sl *ljubiti* : 'loves'.
6. eyebrow : S *bhrū* ; Av *brvat-*; Gk *o-phrus* ; C *brūad*; Gmc *brūn* ; B *bruvis* ; Sl *brŭvĭ* ; Toch A/B *pärwān(e)* .
7. flesh: S *māṃsa*; Gmc *mimz*; B *meisa*; O Prus *mensā*; Sl *meso*; Alb *mish*; Arm *mis*; Toch B *misa.*
8. heel : S *parṣṇi*; Av *pāsna-*; Ht *parsna-*; Gk *pternē*; L *perna*; Gmc *fiersn.*
9. jaw : S *hanu*; Av *zānu-*; Gk *genus*; L *genu-* (back-tooth); C *gen*; Gmc *cin/kin*; B *žan-das*(?); Toch A *śanw-em* (fem dual).
10. knee : S *jānu*; Av *žnu-*; Ht *genu*; Gk *gonu*; L *genu*; Gmc *kniu*; Arm *cun-r*; Toch B *keni-*.
11. nose : S *nas-*; Av *nāh-*; L *nārīs*; Gmc *nasa*; B *nosis*; Sl *nosĭ.*
12. palm (of hand) : S *pṛtha*; Ht *paltana*; Gk *palamē*; L *palma.*
13. reward : S *mīḍha* ; Av *mīžda*; Gk *misthos;* Gmc *mizdō;* Sl *mĭzda.*
14. shoulder : S *aṃsa*; Gk *ōmos*; L *ume-rus*; Gmc *ams*; Arm *us.*
15. tooth, molar : S *jambha* ; Gk *gomphos* ; (B *žam̃bas* 'sharp edge';) Sl *zǫbŭ* ; Toch A/B *kam/keme* .
16. woman : S *gnā/jānī* ; Av *jani-*; Gk *gunē* ; Gmc *cwene/qino* ; Sl *žena* ; Arm *kin* .

Note that there are some 40 stems designating parts of the human body which are the most invariable of all words since people have their bodies in all places and conditions and can hardly confuse eyes with ears or breasts with cheeks. One might dispute 10 or 15 of them but 25 at least are certain cognations.

The question now arises: Why should, say, the Baltic-speakers give up their native non-IE word for 'belly' (assuming it was different) and adopt *vēderas*, changed from the IE stem used by the Greeks (*hoderos*) or the Romans (*uterus*)? Or, why would the Germanic-speakers give up their own non-IE word for 'jaw' and adopt *cin/kin*, changed from the IE stem used by, say, the Celts (*gen*) or the Romans (*genū*)? ... Well, yes we can hypothesize successively at question mark after question mark, but at some point we must stop this easy way of escaping and get down to facts and sound reasoning. Yes, indeed, there may be some few freaky examples of inexplicable borrowing. But in the end the answer is– there is no ostensible reason whatever. Such changes could come about on a large scale only through contact with a superior (and literate) culture or through conquest and coercion. Otherwise, there must be a genetic relation.

(c) The similarities extend to genders, verbs and terminations. I take some random examples. Consider:

i) The forms and endings of the 1st and/or 3rd sing of the non-reduplicating Perfect of the verb 'to know':

S	AV	Ht	Gk	L	Gth	B(OPr)	Sl
veda	***vaēδa***	-	***Foida/e***	***vīdi(t)***	***wait***	***vaīdi-***	***vĕdŭ***

Ht has only a Periphrastic Perfect which is formed with the nom/acc of the neuter participle plus the auxiliary *har-ak-* 'to have' as in *markan har-ak-* 'to have cut'. Cf similar periphrasis in S *vidām as-/kṛ-/bhū-* 'to have known'. This example of Perfect as given in the forms above implies a whole conjugation in existence in all these branches and this precludes borrowing or a limited group.

ii) The same should be said for many more like the following verb 'one is' (3rd pers sing):

S	Av	Ht	Gk	L	Gth	B(Lith)	Sl
asti	-	***ēšzi***	***esti***	***est***	***ist***	***ēsti***	***jestŭ***

iii) Consider the endings for, say, 1st and 2nd plural Pres Active of 'carry'.

S	Gk	L	Gth	Sl
*bharā**mas***	*phero**men***	*feri**mus***	*baira**m***	*bere**mŭ***
*bhara**ta***	*phere**te***	*fer**tis***	*bairi**þ***	*bere**te***

iv) The present participle active of the verb 'to eat'

S	Ht	Gk	L	OE
*ad-**ant**-*	*ada-**anz**-a*	*ed-**ont**-*	*ed-**ent**-*	*et-**end**-e*

This too precludes borrowing or the limited group.

Here all the examples above show a wider spread including Gmc, B and Sl. Thus the notions of borrowing and of the limited group are nullified. The same is demonstrated in the noun example that follow. More on this at the end of this section in paragraph (d).v) Take the sing Gen for 'father' (using 'mother' for Sl); note that in some branches like S this is also Abl.

S	Gk	L	Gth	Sl
pitus/r/ḥ	*patros*	*patris*	*fadars*	*(matere)*

vi) And the sing Gen for 'dog'

S	Ht	Gk	Lth	OIr
śunas	*kūnar*	*kunos*	*śums*	*coin*

vii) Also sing Gen of 'cloud'

S	Av	Ht	Gk	Sl
nabhasas	*nabaŋhas*	*nepisas*	*nepheos (nephous)*	*nebhese*

Note that these examples also exclude chance similarity and the narrow group. Now let us examine oblique cases other than the Gen.

viii) Consider the word for 'name' in sing NVA and Loc and Pl NVA.

	S	L	Gth	Sl
Sing NVA	*nāma*	*nōmen*	*namō*	*imę*
Loc	*nām(a)**ni***	*nōm**ine***	*nam**in***	*im**ene***
Pl NVA	*nāmā(**ni**)*	*nōm**ina***	*nam**na***	*im**ena***

Note that Gk has *onoma* for NVA, *onoma-t-**i*** for Dative (=Loc) and Pl NVA *onoma-t-**a***.

Some of the above examples are taken from Misra (1968), Held (1997), Szemerényi (1996) and Clackson (2007). Here, all these similarities can be due neither to accident nor contact and diffusion. Many more are found in other papers.

It may be argued that other languages (usually reconstructed Proto-this and Proto-that) show certain affinities (nominal or verbal terminations) but these are very few and very tenuous and do not prove borrowing or convergence due to long/repeated

contact. They may be due to accident that is independent native change, or to inherited memory from a distant common past. The similarities in the IE branches have sufficient divergence to show that they are not direct borrowings and sufficient similarity to show that they are devolutions from common original forms. Moreover, intermediate Near Eastern (non-IE) languages do not have these elements.

(d) Some of the features in (b) and (c) may not occur in statistically significant numbers, as Angela Marcantonio has argued (2002, 2005), but this is not a significant objection. It is generally assumed that statistics is a vital constituent of the "scientific" method. It is not realized that the scientific method is not applicable to many areas of human life and that, in any event, it is as much dependent on inspiration and luck on the one hand, and as much liable to blunder on the other, as any other "method" (Pert 1997; Cohen 1985; Beveridge 1968). In any event, in the area under discussion, the proper "scientific" approach would be to ignore rigidity, regularity and uniformity since the linguistic and cultural changes occurred in diverse ways and certainly under no observable law as we shall see in §§6-7.

Then, statistics is a purely quantitative factor that has no bearing on quality or significance - matters which require different criteria[3]. Statistics and numbers here must give way to more important considerations like the improbability of so many similarities resulting from sheer accident or from various diffusion waves. Apart from subjection or migration and the like, it is most improbable that, in normal circumstances, any ancient people would adopt the foreign linguistic features examined in (b) and (c), which in some cases would make speech extremely difficult. Consider the complexities alone in the conjugation of the reduplicating verbs in Greek (in *-mi*) and in Vedic. In historical times the tendency for languages is to become simpler, more streamlined by analogy and more synthetic with prepositions and auxiliary verbs. Consider also the difficulty mentioned in (a) that all such similarities are absent in Near-Eastern languages.

It may be and has been claimed that some languages or groups of them have innovated: e.g. Sanskrit, Greek and Latin (and less Hittite) spent time together and thus developed their highly complex verbal system to the exclusion of Germanic, Baltic etc. Such generalizations are entirely conjectural, unproductive and, since they don't embrace all relevant facts, totally unconvincing. First, no IE language in the historical period is known to develop from a simple to a complex inflective state, as has just been said; the contrary is observed in the cases of (various Afro-Asian languages, too) Vedic, Greek Latin and, even, Old English which devolved into a simpler non-inflected language increasingly using prepositions and auxiliary verbs. Second, there are lexical items and isoglosses that are shared by groups other than Sanskrit, Greek and Latin; thus palatalization separates Sanskrit from Hittite , Greek and Latin and joins it to Slavic and Baltic. Third, certain grammatical features join together Sanskrit Greek and Baltic - like the sigmatic Future; moreover, the reduplicated Perfect is found in Germanic as well as Sanskrit, Greek and Latin but not Baltic. To these should be added nominal inflexional features like neuters of the *nāma/nomen* type (in §4, c, viii) which are shared by Sanskrit, Latin, Gothic and Slavic (but Gk has stem *onoma-t-*). Thus the notion of a group consisting of Sanskrit, Greek and Latin (and Hittite) must be abandoned. Otherwise, we must also wonder whether Slavic, Baltic and Germanic also started innovating just like Sanskrit in the same areas (by sheer coincidence!) and then suddenly stopped. Such claims, therefore, are valueless.

4. Another important type of evidence for a unitary PIE civilization is the presence of identical or very similar **cultural features** in the different branches.

(a) The spread of theonyms is our first consideration. We find a Firegod with the same or cognate name in three branches and nowhere in non-IE languages: thus V *Agni*, Ht *Agnis*, Sl *Ogon* (and variants) - while the stem for 'fire' (but not the theonym) appears as L *ignis*, Lth *ugnis* and Iranian *dašt-aγni-*. The name of another god appears as V *Aryaman*, Mycenaean *Areimene* (and probably later Gk *Arēs*), C *Ariomanus* (in Gaul)/ *Eremon* (in Ireland) and Gmc *Irmin*. The Skygod is V *Dyaus*, Ht *ᴰSius*, Gk *Zeus/Dia-*, L *Ju[s]-piter*,

Gmc *Tīwaz* and South Russian *Divŭ*. One Rain- or Thunder-god is V *Parjanya*, Sl *Perenŭ*, B *Perkunas* and Gmc *Fjorgyn*. Then the Sungod is V *Sūrya*, Gk *Helios*, L *Sol* and B *Saule* – while the stem for 'sun' appears as Gmc *savil/sol*, Welsh *haul* and Sl *solnce* (and variants). The Dawngoddess is V *uṣas*, Gk *Ēōs*, L *Au[s]rora*, Gmc *Eos-tre* (=Spring) – while the stem for 'dawn' appears as B *aušra* and probably C *gwawz* (Kazanas 2006a). Surely all these agreements are not accidental.

(b) Many religious, ritualistic practices are shared by several IE branches. A well-known one is the horse-sacrifice, found in one form or another in India, Greece, Rome, Germania and Ireland (for details see Andersen 1999). But they share also several customs and legal practices (eg prohibition of incest, division into five tribes or demes/regions, etc), apotropaic and purificatory rituals and, of course, divination (Dumézil 1952, 1954; Puhvel 1970; Kazanas 2001 for Greek and Vedic).

(c) Poetics too is a feature common in IE branches. Here we must exclude Hittite poetry which falls wholly within Near-Eastern traditions and has very little or no relation to IE forms and themes. Moreover, since Latin poetry imitates largely Greek prototypes (epic, lyric, drama) and Celtic, Germanic etc are again largely imitative of Latin and Greek forms (Watkins 2001), comparisons here must be made between Greek and Indic materials. In Vedic and Greek poetry we find three types of stock epithet: the Greek ones are all culled from *Iliad* 1 and the Vedic ones from the *Ṛgveda*. (Much of this comes from Kazanas 2001a.)

(i) Vague adjectives like Gk *dios* 'bright, divine', *diogenēs* 'nobly born', *megathumos* 'big-hearted'; etc. Corresponding Vedic ones are *daivya* 'bright, divine' (*RV* 1.35.5; 2.33.7), *ugra* 'fierce, mighty' (2.33.9; 10.34.8), *ṛtavan-* 'holy, observing order' (2.35.6; 7.61.2); etc.

(ii) Epithets denoting a specific feature that could be used of anyone but are applied only to a hero or divinity: e.g. Gk Hera 'of white hands' *leukolenos*, Athena 'of blue/grey eyes' *glaukopis*, Acheans 'of fine greaves' *euknēmis*; V *somapā* 'somadrinker' could be used of any god but is applied only to Indra (2.12.13; etc), *jalāṣa* 'cooling' used of Rudra (2.33.7, 7.35.6), *jātavedas* 'all-knowing' of Agni (1.44.1; 4.3.8), etc.

(iii) Epithets used of one figure (hero or deity) and denoting a specific feature thereof: Gk *asteropētēs* 'who throws the bolt' for Zeus, *hekēbolos* 'aim-attainer, farshooting' for Apollo, *polumēchanos* 'of many devices' for Odusseus; V *gṛhapati* 'lord of house' for Agni (1.45.1), *vajrin* 'he of the bolt' for Indra (7.49.1), *urugāya* 'far-going' for Viṣṇu (1.154.1) etc.

(iv) Then there are items of a common lexical stock (again in the *Iliad* and the *RV*): V *śravas/śruta* = Gk *kleos/kluto* 'fame(d)'; V *uru* = Gk *euru* 'wide'; V *āśu* = Gk *ōku* 'swift'; V *patnī* = Gk *potnia* 'reverend lady'; etc, etc.

(v) In Vedic poetry we meet strict metre like Anuṣṭubh (4 lines of 8 syllables) with an iambic cadence in the second and fourth unit, or *Gāyatrī* (3 lines of 8 syllables) but also within these strict metrical lines rich alliteration. In Greek and Latin poetry we find different metrical units (iambus, anapaest, hexameter etc) while in Germanic we find alliterative lines with loose metre.

All these features, except for rare recurring epithets, are not found in ancient Near-Eastern literature.

(d) Common incidents, motifs, themes in religion (=mythology). A most interesting motif is that of **the goddess who becomes a mare** and a pursuing god turns into a stallion and mounts her: from this union comes a twin-pair or some other marvellous creature. In the Vedic tradition we have Sungod Vivasvat who becomes a stallion to mount his run-away bride Saraṇyū who had turned into a mare and later gave birth to the twin Horsegods Aśvins (*RV* 10.17.1-2; *Bṛhaddevatā* 6.162ff). In Greece (Pausanias VIII.25.5), Poseidon chases Dēmētra Erinus (= V *Saraṇyū*?) then both take on equine form and later Dēmētra gives birth to the beautiful horse Areion and a daughter Despoina. Finally in Norse legends (*Edda*, 35-6), Loki, the god of tricks and transformations, becomes a mare to attract from work the giant-mason's stallion, Svadilfari, then gives birth to wondrous Sleipnir, the eightlegged swiftest racer in the world, given as gift to Odin (see Kazanas 2005a: §4).

Another motif is that of the heroes or **divine youths who rescue** the Sunmaiden from a dangerous situation or become her companions. In India it is the Aśvins (saviours of men from tempests and other calamities) who accompany Sūryā the Sun-maiden, even to her wedding. In Greece it is the *Dioskouroi* 'Zeus's lads', Castor and Pollux (also saviours and expert horsemen) who rescue their abducted sister Helenē. Among the Lithuanians it is the *Dievo Sūnelai* 'the [sky-] god's lads' who rescue and escort the Sunmaiden *saules dukterys.*

A third example is the unusual motif of **the thigh-born child.** In Greece we meet the mythologem of Dionusos being born from the thigh of Zeus. In one version Semelē asked Zeus to appear in all his brilliant glory while she was six-months pregnant with his child; he did so, blazing away with his lightning and Semelē was burnt up. But then Hermes rushed and saved the embryo and stitched it within Zeus's thigh. The earliest attestation is in Euripides's *Bachai* 88-100 (*kata mērōi de kalupsas* 'having covered [the embryo] within the thigh' 97) which means c420. In the Indian epic *Mahābhārata* Bk I, 169-71 there is the story of a radiant brahmin lady who fled from some cruel warriors carrying her child in her thigh; they found her but then the child issued out of the thigh blazing like the sun and blinded them. Here it may be claimed that the Indian tale is late and perhaps borrowed from Greece. But in the much earlier *Jaiminīya Brāhmaṇa* (III, 199), which would be 5th c. at the latest, the child Kutsa is born from Indra's thigh: no details are given here, but Indra was, like Zeus, the storm-and-lightning god (Kazanas 2004c: 46-49).

Let us take a fourth and final motif – **the severed head.** In Norse legendry, Odin preserves wise god Mimir's head (cut off by the Vanir) to consult it in times of danger and doubt (Davidson 1981:146). In the Welsh tradition the family of Lyr preserve Bendigeidfran's head (MacCana 1983: 78). In Greek myths Bellerofon holds Medusa's head which still has the power to turn the onlooker into stone; then, Orpheus's head, after the Maenads tore him to pieces, floated down the river Hebros still lamenting and at Lesbos was installed as a shrine of prophecy. In India, in the *RV*, the Aśvins substituted sage Dadhyañc's head with that of a horse to obtain secret knowledge and when Indra cut off the horse-head they reinstated the original. (See Macdonell 1898:141-2.)

Many more such motifs and parallels will be found in Kazanas 2001b, 2004c, 2005a and 2006a, all of which appear, as expected, with divergences in the east (=India) and in the west (Europe), but do not appear at all in intermediate Near East; these also suggest a common origin.

5. The aspects examined in §§ 3-4 indicate clearly that there is sufficient evidence to justify the claim for a unified PIE civilization which had a definite location somewhere in Eurasia before 6000 BC. It is generally assumed that 'civilization' means tools, large buildings, statues and paintings, ornaments, vehicles, weapons of war and other material artefacts. But civilization can flourish without advanced technology and artefacts, such as we find in Ancient Mesopotamia and Egypt: it can develop and thrive on elementary technology and very simple agriculture for a very long period of centuries and millennia. The difficulty will be that without the material evidence we don't know that there had been an advanced civilization. R. Rudgley cites a modern scholar, Prof Yoshinory Yasuda who found "a marvellous principle" for civilization which is "respect for and co-existence with nature" and added: "Civilization begins to appear when a workable system for living, that is a proper relationship between men and nature, is established in accord with the features of a given region". Yasuda wrote this in regard to the Jōmon culture in Japan beginning c 10 000 BC (Rudgley 1998: 31-33). No doubt there are many more definitions. One definition that is interesting because very unusual in directing our attention inwards is by A. West who wrote about Ancient Egypt: "In a civilization, men are concerned with the quality of the *inner* life rather than with the *conditions* of day to day existence... a concern... to master greed, ambition, envy" (1993: 6-7). Then, Plato's first ideal society is a community with simple agriculture, animal husbandry, essential crafts and trade (exchange), feeding on barley-bread and bulbs, drinking wine in moderation and singing hymns to the gods (*Republic* 370 C ff). The PIE civilization may have been of this nature with a

"proper relationship between men and nature" and perhaps "a concern to master greed, ambition, envy" as their highest aspiration.

The evidence we have examined indicates clearly a unitary PIE civilization. The same evidence has sufficient diversity to show just as clearly that this civilization, wherever it was located (and this could have been an extensive area), lost its unity as, at a first stage perhaps, some sections of its population in different regions began to lose their pristine character deviating from the original norm and developing distinct linguistic and cultural peculiarities. At a second stage, again, different sections, large or small, successively moved away to new areas and eventually, settled to their respective historical habitats - Celts, Germans, Greeks etc. In this they influenced and were influenced by other cultural groups they encountered on the way or at the locus of settlement. Some lost many aspects of their original PIE culture, others lost most (at different indeterminate periods) and still others retained only meagre linguistic elements. Hence the considerable diverse cultural groups which, mainly through linguistic studies, have come to be known as the IE branches.

6a. Proto-Indoeuropean ? From the linguistic fossils in the IE branches modern scholars have reconstructed what they claim to be the forms of words in the PIE language. They even call it a "science"; but it all seems wishful thinking. I have argued elsewhere (Kazanas 2002b, 2003, 2004b) that reconstructions of Protolanguages are unreliable because they are conjectural and unverifiable: even if they happen to be right, we shall never be certain of this. O. Szemerényi admits that reconstructions are used to facilitate comparisons, using one word instead of many IE variants, and cites Hermann's statement that "complete forms (e.g. *deiwos [=S *deva-s*]) cannot be reconstructed at all, only single sounds, and even these are meant as approximation only" (Szemerényi 1996: 33; my square brackets). Twenty years earlier Burrow had said much the same: "in the case of Indo-European it is certain that there was no such unitary language which can be reached by means of comparison... the Indo-European that we can reach by this means was already deeply

split up into a series of varying dialects" (1973:11). More recently, exhibiting scepticism like mine, X. Tremblay writes (of various IE branches but mainly Iranian): "la grammaire comparée est en réalité *radicalement* incapable de discriminer entre parenté divergente ... et parenté convergente" (2005: 63). This does not stop comparativists from reconstructing PIE forms and evince considerable faith in their 'scientific' reconstructions. And here we meet a serious instance of a conflict/divergence between thinking and acting that indicates great confusion.[4]

Now, let me take a simple example as starting point. Although I often use Rix's *Lexikon der indogermanische verben...* (e.g. §17, below), I cannot but agree with A. Marcantonio's critique of his methodology in taking often only 2 similar stems which, indeed, may be loans or may be devolutes of one stem in two closely related branches, e.g. Baltic and Slavic or Germanic (see her paper in this volume). The examples I gave earlier (§3,b) all have three or more cognations and encompass eastern and western branches. So, to start with, many stems may not be Proto-Indo-European and all such reconstructions are obviously pointless and utterly misleading.

Linguistic changes (vocabulary, accidence, spelling etc) are not subject to universal laws. The way English has changed from 1100 CE is quite different from the way French or Greek changed, even though some aspects are general and common (loss of inflexion, increasing use of auxiliaries etc). Now, while certain general phenomena have some regularity and invariability and thus may be said to approximate the nature of "law" within documented and defined periods of changes in any one language or in relationships within a group of languages (like the IE), nonetheless all such phenomena have exceptions and, in any case, are specific to the particular period of the language or group of languages under examination. Marcantonio (2005 and in this volume) rightly criticizes many aspects of these so-called "laws"[5.] The changes in Vedic, Greek, Germanic and other languages differ enormously among themselves: eg the various forms of 'be/become' as in S √*bhū* (>*bhava-*), Gk *phuomai*, L *fui*, C *buith*, Gmc *be-* etc. Then again, Gk plosive *ph* became a fricative *f* as is the Italic *f*.

How or why did the original initial consonant – whatever it was – change into these sounds? Linguists don't know. One comes across various hypotheses but there is no sure knowledge – because there is no documentation. The changes occurred in particular, if not precisely determined places, periods and peoples and stopped there. If linguists don't know how or why these changes occurred, then they most assuredly cannot know from existing material what the original forms were. It may be legitimate to make guesses and postulate certain forms (for the sake of convenience). But to proceed then to regard these entirely hypothetical forms as factual and use them in further comparisons with other hypothetical forms or build upon them structures and employ them as arguments in discussions about historical events seems to me wholly unacceptable.

6b. Consider a different example. In Greece, in the late 1970's an Act of Parliament imposed reforms in the spelling, accent and accidence of Modern Greek. Nobody could have predicted these changes 100 years earlier just as nobody could have predicted at c900 CE the changes in English that would come with the Norman Conquest or later, after Chaucer. All such changes have no regularity, constancy or universality and are of the nature of accident – not law – due to social factors. If we can have no reliable predictions about future developments we can have no reliable reconstructions of forms in past periods before documentation: so also most emphatically Hock 1991. Reconstructions are futile, wasteful exercises.

Another obvious inconsistency is the reconstruction of three dorsals which are thought by many to be unnecessary (Szemerényi 1996; Sihler 1995; Lehman 1993; Speirs 1978) and one of which, the labio-velar family, is unpronounceable! How on earth can anyone pronounce as one unitary sound k^w or g^wh which contain both a guttural/velar and a labial element? One simply can't, however neat it looks on paper. To state that labio-velars exist in Latin (and other languages) is an assertion of no value. Certainly, we have L *qu-a-(lis/* etc) and in English *qua*-(lity) and Italian *quando* etc. But in every case we have three separate,

successive sounds: velar *q* (pronounced **k**), a glide vowel **-u-** and a vowel; in no case do we have a single, unitary consonant. In fact, the phonemes **qu**- do not differ from L **cu**-(lpo). If it is claimed that L **qu** is pronounced **q-v-** then, again we have two distinct sounds – the velar **q/k** and the labial semivowel **-v-**. Sanskrit also has *kva, dva, śva* etc: these can be pronounced only as separate successive sounds not as a single one. Attempts at describing how these hypothetical phonemes are pronounced (e.g. Clackson 2007: 49-51) are theoretical and of no practical value. Please, try yourselves! Also to pronounce the monstrosity *dhg^whec 'to perish'! (For other unpronounceable examples see, for instance, Watkins 2000, 2001; Clackson 2007.)

A final point about laryngeals, which are wholly unattested except in Hittite (a language whose IE character is much decayed, being smothered by other neighbours rich in laryngeals). Let us consider the example of H_2 (or h_2 or $ə_2$). This appears in PIE *$dhugH_2ter$ 'daughter' (Fortson 2004: 204): appears as *a* in Gk *thugatēr*, but as *i* in S *duhi*tar. However, Av *duɣðar* (Hale 2004: 748) or *duxtar* (Fortson, 204) has neither *a* nor *i*. So what was the form in Proto-Indo-Iranian?... Not known. Old Avestan has *pta* for 'father' but later *patar* and *pitar* (Mayerhofer *KEWA*, vol 2, 277); this is S *pitṝ*, Gk *patēr* and L *pater* (Fortson, 23, 276) all allegedly from PIE *pH_2ter. But, again, what was the from in Proto-Indo-Iranian?...Unknown. First of all consider that unlike S (which has many cognates from √*duh*), Av *duɣðar* and Gk *thugater* stand isolated without related stems in their languages. Then, as M. Hale observed, the *i* was not an invariable feature of Proto-Indo-Iranian (2004: 748). The cognates for 'father' expose yet another inconsistency. L has also *Ju[s]pitar* with *i* as well as *pater* in the selfsame phonetic environment and B. Fortson offers no explanation at all (2004: 23, 33, 253, 261; also *Marspitār* 276, 406). A further difficulty sprouts out from S *pitr*. We mentioned already that Av has *pta, pitar* and *patar*, despite the selfsame phonetic environment. But according to IEnists S should have **phitṛ*! Because according to the IE "reconstruction"-system, the laryngeal H_2 becomes *i* in S but also aspirates the previous

consonants. Thus PIE *stH_2to > S ***sthita*** 'one who has stood' and PIE *$pletH_r$- > S ***prathiman*** 'width'. However, *pitṛ* has no aspirate *ph*! Those disparate phenomena show most flagrantly that these IE "reconstructions" and "phonetic laws" are anything but satisfactory.

7a. A most interesting aspect of Indoeuropean Linguistics is the root and the vowel gradation or *ablaut* system (or apophonie). Indoeuropeanists adopted an hypothetical five-grade ablaut from Greek. Now, the only language that has roots is Sanskrit. To begin with, the word 'root' does not strictly translate the S *dhātu* 'layer, element, constituent, seed-form'; nor can a "root" exist as an independent and generative element of a plant as a seed can. But putting this aside, only Sanskrit has roots and a proper vowel gradation. **All other IE branches have stems, not roots as such.** Like every other modern IE branch English has no actual working concept of root. (Of course 'act' in English can function as a root and generate '***act***-ed, ***act***-ing, ***act***-ion, ***act***-ivate, in-***act***-ively' etc, but 'act' itself comes ultimately from Latin.) Greek had verbs, and scholars say that nouns derive from the verb-stem: e.g. *che-ō* > *che-u-ma* 'a flow/stream'; *cho-ē* 'pouring, libation', *cho-a-nē* 'melting pot'; *chu-ma* 'the fluid', *chu-s-is* 'shedding', *chu-tra* 'earthen pot'; etc. Even if we took *che-* as the root, it is difficult to see how this develops into *cheu-*, *cho-* and then *chu-*! One realizes how inconsistent Greek is when one considers two similar verbs: *deō* 'bind' > *de-ma* 'band, rope', *de-s-is* 'the binding together', *de-s-mos* 'bond', *(dia-)dē-ma* 'ribbon round hair' – but no *deu-*, *do-* and *du-* ; *pne-ō* 'blow, breathe' > *pne-u-ma* 'blast of air' (later 'spirit'), *pne-u-s-/pno-ē/pnoi-a* 'blast, breeze, breath' – showing unexpected *pnoi-* but not *pnu-*! If one examined other similar verbs (*bdeō*, *zeō*, *keō*, *xeō*, *neō* etc) one would find even more bizarre changes in the stem. There is no regularity; moreover the vowels change from palatal *e* to labial *u/o* etc without rhyme or reason.

Sanskrit has three gradations in the development of the root-stem: e.g. √*cit* 'being conscious of' > *cet-as* 'mind, intelligence' or *cet-a-ti* 'he/she realizes', *a-cait* 'realized' (aor), *cait-anya* 'consciousness' etc. ˇ always changes to *e* and *ai*, never to *a* or *u/o*.

Similarly radical *u* →→ *o* →→ *au* and *ṛ* →→ *ār* →→ *ār*. Now, *ṛ* sometimes will give *ra/ri/ru* but will never become *i/e* or *u/o*. Thus there is the basic grade of the simple vowels *a, i, u, ṛ, ḷ* (though some roots have a 'developed' vowel), the strong (*guṇa*) grade *a* (same), *e, o, ar, al* and the fully developed one (*vṛddhi*) *ā, ai, au, ār, āl.* This triple gradation has its equivalent in Vedic cosmology and philosophy where we find three main world-levels – heaven, midair and earth (*svar*, *antarikṣa* and *pṛthivī*) – and causal or natural, subtle or mental and gross or material. As nouns and verbs are generated from the root , the radical vowel changes according to constant regulations (except, as was said, in the case of *ṛ* which is somewhat unstable). This process is absent from other IE branches. (And, as we see in Greek, it is utterly confused. Modern studies since, say J. Kurylowicz (1956), O. Szemerényi (1972) *et al*, rationalize but also tacitly acknowledge this fact.)

After Brugmann had turned against Sanskrit and favoured Greek (1912-16), comparativists adopted a five-grade ablaut from Greek *patēr* 'father' and *eupatōr* 'good father' in incredible disorder (Szemerényi, p84):

i) *pa-ter-a* (acc sing) where *-ter-* shows ***e*** as the basic grade (to be distinguished from S *e* which is long and second grade).

ii) *eu-pa-tor-a* (acc sing) where *–tor-* shows the **o**- grade (again to be distinguished from S *o* which is long). But note that, unlike the Sanskrit series, here we have a different noun declension and sound-family: *e* is palatal and *o* labial!

iii) *pa-tr-os* (gen sing) where *–tr-* shows the zero or nil grade! But here we have in fact syncopation or loss of vowel (*lopa* in S).

iv) *pa-tēr* (nom sing) where *–tēr* shows the long-vowel grade (**ē**).

v) *eu-pa-tōr* (nom sing) where *-tōr* shows the long ō grade (**ō**-mega).

Obviously there is neither rhyme nor reason in all this. More importantly, the ablaut occurs in the affix *-t-r-* not the root stem. This series, if it can be considered such, might have some validity if it occurred in one noun or one verb and if it followed some principle(s) that governed the same changes in many other nouns or verbs. Szemerényi gives instead the following examples stating that "Very often only full grade [i.e. the vowel *e*], o-grade and zero grade are attested" (p84):

a) *leip-ō* (pres) *le-loip-a* (perf), *e-lip-on* (aor) 'leave'
b) *derk-omai* (pres), *de-dork-a* (perf), *e-dr-akon* (aor) 'perceive'.
c) *penthos* (neut nom sing), *pe-ponth-a* (perf), *e-path-on* (aor) 'grieve'.

Here one cannot but wonder at this astonishing presentation, which merely increases the disorder. Here we have more bizarre phenomena. First, the vowel-change occurs in the stem, not in affixes or terminations, as above. Then *leip-* and *derk-* are verbs but *penthos* in (c) is a noun! Then the zero grade is not syncopation or disappearance of the vowel as in (iii) above, but a different vowel ***–i*** in (a) and ***a*** in (c). Another difficulty is the dipthong *ei* in (a); for we find also *eu-pa-**teir**-a* (nom sing fem 'she of a noble father') and we now do not know whether this is full grade as with *leip-ō*, or a sixth grade according to the five-grade presentation above[6].

But what principles govern these so-called vowel-grades in Greek? In fact there are no principles such as we have in Sanskrit! Clackson writes at length about the significance of the root but does not give a single example in PIE or any branch of a root generating primary and secondary derivatives as in Sanskrit (2007:65-9, 187-191). This issue and others like the obvious problems of reconstructed sounds, especially consonants, I have discussed at great length elsewhere (Kazanas 2004b) and we need not therefore spend more time with them. A quotation from Burrow should suffice here: "Chiefly owing to its antiquity the Sanskrit language is more readily analysable and its roots more

easily separable from other accretionary elements *than any other IE language* (Burrow 1973: 123, 289, my emphasis).

7b. By way of conclusion, let me repeat that PIE cannot be reconstructed at all from the available data in the known IE branches. The comparative efforts fail as a science miserably, because they employ far too many "laws" which at the same time have far too many exceptions and discrepancies. One of the worst aspects is the reconstruction of words that, as we saw, look very nice on paper but cannot be pronounced – like **nekʷts* 'night' or **dhĝhm̥ma* 'on the ground' etc. The positive sciences have as their final testing ground Maths and the reality of the material world. What corresponds to this in the linguistic/phonetic field is sound, *pronounceable sound*! If the lexemes we produce are unpronounceable, then our system is wrong.

A linguist not involved in IE comparativism, R. Nixon (Professor of Linguistics at Australian National University) points out sharply (1997: 46) that R. Hall reconstructed the Original for the Romance languages (Italian, French, Spanish etc) but this was only a partial approximation to Latin, though some elements were correct. F. Southworth made a similar study of the various Indoaryan languages and his result also was only a partial approximation to Sanskrit. Nixon points out that each scholar knew the respective original language. Nixon's implication is that not much of the PIE reconstruction could be correct when the IEnists do not and cannot know the PIE since there are no documents in the PIE languages. Enough said on this subject.

Indo-Aryan

8. Having held and taught for more than 18 years, but without investigating, the received doctrine that the IE branches dispersed from the South Russian or Pontic Steppe (as per Mallory 1997, 1989; Gimbutas 1985, 1970; and others), and that the IAs had entered Saptasindhu c1500, I began to examine these mainstream notions thoroughly and c1997 abandoned them. I decided that no evidence of any kind supported them; on the contrary, the evidence showed that by 1500 the IAs were wholly indigenous and

that the elusive IE homeland was very probably Saptasindhu and the adjacent area – the Land of Seven Rivers in what is today N-W India and Pakistan; this area could well have extended as far northwest as the Steppe.

Apart from the (recent) genetic studies, which at that time were not so well-known nor so secure (see §19, below) the decisive evidence for me now is the antiquity of Sanskrit indicated by its inner coherence and its preservation of apparently original PIE linguistic features (like the dhātu, five families of phonemes, etc) and cultural elements (e.g. §4). The Vedic language as seen in the *RV* alone, despite much obvious attrition and several innovations, has preserved many more features from the putative PIE language and wider culture. This was due to its well attested and incomparable system of oral tradition (Kazanas 2002, 2003, 2004b, 2006a) which preserved the ancient texts fairly intact (*RV*, *AV*, etc) and continued even into the 20th c. An oral tradition of this kind cannot be maintained by a people on the move for decades if not centuries over many thousands of miles, as the AIT proposes. Such a tradition could be preserved only by a sedentary people where the older generation would have the necessary leisure to pass the communal lore to the younger one[7].

9. I was not of course the first to come up with such a view. On the strength of Sanskrit many European scholars in the early 19th c. thought India was the original homeland (Mallory 1973:26-9). Even after scholars rejected India in the later 19th c and began to adduce different urheimats from the Baltic to the Balkans rewriting Indian (proto-)history, there was a succession of Indian scholars, mainly, who maintained the indigenist view in one or other form: Rao 1880; Shri Aurobindo 1914; Dhar 1930; A. Das 1971 et al. Indeed, in recent years also many publications advocated indigenism: Sethna 1992; Elst 1993, 1999; Frawley 1994 and with Rajaram 1997; Feuerstein et al 1995; and others[8]. In fact, more and more scholars in the West have re-examined the issue and rejected the mainstream view advocating instead a movement Out of India into Europe: Schildman 1994, 1998; Elst 1999; Friedrich 2003, 2004; Danino 2007. Here we must add the fact that since

1984 there is full consensus among archaeologist specialists in the area that the ISC has unbroken continuity (Lal & Gupta 1984; Allchin 1997; Kenoyer 1999; Shaffer & Lichtenstein 1999; McIntosh 2002; Possehl 2003). Thus the mainstream strident cries of warning about "fundamentalists, nationalists, revisionists" in India and their colleagues in the West (Witzel 2003; Huld 2002; Kuzmina 2002) are seen to be either hollow slogans or echoes of those slogans; for neither the indigenist scholars mentioned above nor the archaeologists specializing in the ISC and insisting on its unbroken continuity down to the 6th century belong to this (non-existent) band of bogies. I benefited from all of them.

Nonetheless, the AIT persists. Several linguistic arguments have been adduced to prop up this fanciful theory one of them being the (loan-)words in the *RV* language which were allegedly borrowed from other languages met by the IAs on their way to and at Saptasindhu. A closely related argument is the alleged common Indo-Iranian period during this journey to Saptasindhu[9]. Good examples of a combination of these two arguments are found in M. Witzel 2001, 2005 and F. Kuiper 1991. But that 4% of the Rigvedic vocabulary might consist of loanwords (as these controversial studies claim), or even 6% or more, surely should not be surprising since the IAs did not live in a vacuum and there must have been peoples with different languages around them. However, this pet linguistic subject of Witzel's has been attacked by R.P. Das (1995 general critique), by Krishnamurti (2003: 37-8 mainly against Dravidian cognations) and Kazanas (2002b)[10]. In any case, no amount of arguing along these lines gives of itself any actual dates even approximately and the direction of movement can more reasonably be the reverse. There is B. Sergent who, although not at all an indigenist, argued in favour of a movement from Central Asia north-westward (1997). Also, Johanna Nichols, an eminent Indo-European linguist, found on purely linguistic grounds that "the locus of the IE spread was ... somewhere in the vicinity of ancient Bactria" (1997: 113) – which is only a stone's throw away from Saptasindhu. Linguists will continue to disagree violently among themselves about the original locus (from the Baltic to the Balkans and to Bactria),

about the direction of the movement and about the date: some give a late period somewhere c 3000, others an early one, in the 7th millennium (Misra 1992) - and all these disagreements on exactly the same linguistic data! I ignore all such linguistic arguments and conclusions since they are so contradictory. I ignore also the results from the reappearence in Gray & Atkinson (2003) of the discredited method of glottochronology. The situation is in no way improved by claims that "I am right' and allegations that "the others are wrong". We must bear in mind also that all such arguments involve "reconstructed" proto-languages which are sheer conjectural affairs and therefore inadmissible in any impartial court of law.

10. Horses and chariots. One of the best known comparative linguists of our times, Hans Hock, pointed out repeatedly (1999, 1996) that several dialects spoken in N-W India travelled north-westward at different times in the Common Era, the most notable one being that of the Gypsies in the early centuries CE (see also Fraser 1995). Hock stated that Indoaryan indigenism and even the Out of India Theory could theoretically be possible, provided the distribution of the IE isoglosses was accepted and the Vedic language was seen to be a devolved form of Proto-Indo-European - facts which I certainly accept fully. But he opted for the mainstream doctrine (=the AIT out of the Pontic Steppe) because of the archaeological evidence of horses and chariots arriving in Saptasidhu in bulk after 1700 BCE (1999:13). Thus he supports the spread of the Kurgan culture (as do also Allchin B. and R., Gimbutas, Kuzmina, Mallory, Parpola, Witzel et al), even though the cranio-skeletal evidence for this diffusion is very limited according to J. Day (2001). Hock also mentions that it is simpler to have one movement into India rather than many out of it to account for the diversity and chronologies of the spread of the various IE branches. But this is rather self-contradictory since, as he wrote (1996, 1999), the Indoaryan language left India north-westwards three times in the Common Era and generated new dialects and the Gypsies did reach the north-west extremity of Europe into England; so Indoaryan could have done the same in the misty past BCE. Besides, the ancient IAs could, in the Out of

India scenario, have left in a sizeable mass, stayed in a location in Baluchistan/Afghanistan (or even the Steppe) for some time, then, successively, in groups, moved away in diverse directions. We need not make a problem out of this rather simple affair. (But I have argued in other papers that we must not preclude the possibility of a broad IE continuum from the Kurgan steppe down to Saptasindhu at a very early period from which, within which and into which occurred various undetermined movements.)

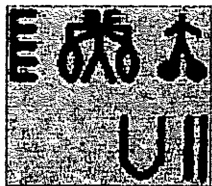
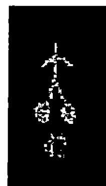
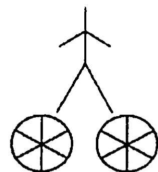
On the matter of horses and chariots, Hock was misinformed, I like to think. The evidence of horses and chariots proves the very opposite if examined with a modicum of attention. As one authority put it, "the horse was widely domesticated and used in India during the third millennium BCE over most of the areas covered by the Indus-Sarasvati [or Harappan] Civilization" (Pande 1999: 344; the square brackets original). I myself collected much evidence for horse-finds in different ISC (=Indus-Sarasvati Civilization/Culture) sites in two papers in 2003, §10 (pp 196-7) and 2002 (p 305): thus one may further consult Kochhar (2000 pp 186, 192), Chakrabarti (1999), Lal (1997), Sharma (1996 p 17), Thomas et al (1995), Dhavalikar (1995: 116-7), Allchin & Joshi (1995: 95), Rao (1991) and, necessarily, Sharma (1980: 110ff). M. Danino's paper 'The horse and the Aryan Debate' is a more recent addition to this literature (2006c). These publications show abundantly that the true horse, *Equus ferus* and *Equus cabalus*, was present in its wild state in India from c 17 000 and there is much evidence now for the domesticated horse in the Mature Harappan period. Moreover there are several terracota figurines of this animal from different ISC sites as reported by Lal (2002) and Tharpar & Mughal (1994).

As for the chariot, S. Piggott established the presence of a sophisticated type of vehicle with "one or two pairs of wheels with their axles... from the Rhine to the Indus by around 3000" (1992: 18).

Archaeological evidence does not consist only of the actual remains of buildings, weapons, tools, chariots etc. Pictures, reliefs, toys and figurines of these things are also evidence. Many years ago H. K. Sankalia had pointed out that the six-spoked wheel

appears on seals and signs of the alphabet (1974: 363). S. R. Rao found at Lothal "terracotta wheels ... with diagonal lines suggesting spokes" (1973: 124). This representational practice seems to have been widespread, for S. Piggott mentions similarly marked wheels found in the Karpathian Basin from the Earlier Bronze Age (1983: 91-92). In his recent study, Lal presents four terracotta wheels (from Mature Harappan sites Banawali, Kalibangan, Rakhigarhi) with spokes painted on (2002:74, Figs 3.28ff). The Harappans had the technology for making spoke-wheels (Kazanas 1999:33; Basham 1954:21).

Finally, it was A. Parpola, an inveterate adherent of the AIT, who identified the figure on Harappan seal No 3357 as representing simplistically a man with outstretched arms standing on two six-spoked wheels (of a chariot) realizing that this was "a later invention of the Aryans' (1969: 24). The later notion that it is a potter on two wheels is obviously far-fetched and, in any case, the wheel is spoked (Sethna 1992)! L. S. Rao has recently presented many more finds of models of terracotta toy wheels with spokes from Harappan locations in *Purātattva*, vol 36 (pp 59-66), 2006.

Spoked Wheel

11. The question of course is whether there were spoked wheels before 3000 since the *RV* uses the word *ara* which is usually translated as 'spoke' (1.32.15; 5.58.5; etc). We don't know. Some Indian scholars approach the issue strategically and say that Rigvedic hymns with *ara* are later intrusions from the second millennium when the spoked wheel was quite common. This is

possible of course, since we know that some hymns are later interpolations. However, there could have been spoked wheels as we know them even at c3200 and before. But then again, *ara* need not mean 'spoke'. It could mean a section of the (solid) wheel. After all, different IE branches have a different stem for spoke (Gk *aktis*, *knēmē*; L *radius*; Gmc *speca/speihha*, etc) which suggests that spokes were developed after the dispersal (in the 6th or 8th millennium, or whichever). So *ara* could well have meant something other than spoke and only later acquired the meaning of 'spoke', as we know it now. The spoked wheel poses, in fact, no problem for dating the *RV*. There are other more clearcut types of evidence which are examined below.

The usual objection is that there are no chariot remains from that period. Here, there is the double assumption that *RV* chariots are necessarily like near-eastern ones of, say, 1500, or the Androvo c1900, and that there should be remains. But why assume that the *RV* chariots are like the near-eastern ones? The only real-life vehicle in a race that we know is in 10.102 and this is pulled by oxen not horses (Kazanas 2002) while 1.123.1 alludes to a 'broad chariot' *pṛthú rathaḥ* ! Elsewhere in the *RV* the chariot *ratha* is described as being *bṛhat* 'tall, big' (6.61.13), also *variṣṭha...vandhura* 'widest... seated' (6.47.9), *trivandhura* 'three seated' (1.41.2; 118.2; 7.71.4; 8.22.5; etc) and once even *aṣṭavandhura* 'eight-seated' (10.53.7)! These are all **rathas** and hardly like the near-eastern battle-chariots. Then again, there are no remains of chariots in India from 1500 or 1000 or 700. So, on the strict AIT thinking, we should say no chariots were brought in by the alleged IE entrants either.

12. Harappan features absent in the *RV*. The Harappan or Indus-Sarasvati Culture has certain characteristics which help to define its uniqueness. A number of these features are absent from the *RV* and this absence indicates that the *RV* is pre-Harappan. Arguments *e silentio* are not decisive since absence of evidence is not always evidence of real absence . But in this case the features are far too many. (Some of these were noted by Sethna, 1992.)

a) *iṣṭakā* 'brick'. The *RV* mentions as building materials metal, stone, mud and wood but not 'brick', which was the basic material in Harappan constructions. This is found in post-Rigvedic texts: the word *iṣṭakā* is not in the *RV*.

b) Urbanization is wholly absent in the *RV*. There certainly was "nomad pastoralism" as mainstreamers emphasize repeatedly but there were also agricultural settlements (a fact which mainstreamers underplay or do not mention). The hymn to Kṣetrapati 'Lord of the Field/Soil' (4.57) alone should suffice but also the girl Apālā refers to her father's *urvarā* 'fertile field' (8.91.5); then there are many cultivation implements *khanitra* 'shovel', *lāṅgala/sīra* 'plough', *sṛṇī* 'sickle', etc. Moreover, there is weaving with loom, shuttle, warp and woof (RV 1.134.4; 1.3.6; etc, etc) and metallurgy with smithies of sorts (4.2.17; 5.9.5; etc). Such activities imply settlement.

Some writers think the Rigvedic and Harappan cultures converge (Gupta 2005, Bisht 1999, Lal 1998, Singh 1995). As evidence is cited the word *pur-* which denotes 'city, citadel, fort, town' as its Greek and Baltic cognates '*polis*' and '*pil(i)s*' do. This is a very general misconception. In the *RV pur* never means anything other than an occult, magical, esoteric defence or stronghold which is not created nor ever destroyed by humans (Kazanas 2003: 224-5 and 2006b passim.)[11]. The ISC cities had regular blocks, large buildings, also domestic and urban water-supply (McIntosh 2001: 100-101): the *RV* knows nothing of all these. There are references to *oka, gṛha, dama*, *dhāma* etc, all of which can denote any type of 'home/house' (made of wood and mud), or the thousand-pillared mansion of kings Mitra and Varuṇa in the sky (probably suggested by sunrays streaming down through clouds 2.41.5; 5.62.6); these most certainly do not indicate any urbanization: neither brick- nor stone-walls are mentioned nor other features as in the ISC towns.

The words for 'council' *sabhā* and *samiti* are also cited but, surely, any community can have a council of elders without urbanization. Allusions in the *RV* to chiefs/kings *rājā* and overlords/emperors *samrāṭ* also do not show urbanization since such offices can just as easily exist in rural communities. (The Red

Indians in North America, nomadic and rural tribes, had local chiefs and overlords.) Pathways and/or roads (*path-*) also have been mentioned as crossing or branching out, but these too can be just as easily seen in a rural context. (For all these claims see Singh 1995; Bisht 1999 and Lal 2002, 2005.)

c) Fixed altars or hearths are unknown in the *RV* but common in the ISC cities. The Rigvedic altar is a shallow bed dug in the ground and covered with grass (e.g. *RV* 5.11.2, 7.43.2-3; Parpola 1988: 225). Fixed brick-altars are very common in post-Rigvedic texts.

d) Many cities were abandoned and fell to ruination after 1900 BC when the Harappans began to move eastward because of the drying up of the Sarasvatī and of the more general desiccation due to tectonic disturbances and climatic changes. The *RV* knows nothing of such ruins even though, according to the AIT, the IAs moved through these regions c 1700-1500 (in small waves, settled there, in the midst of deserts, and wrote the hymns which praise the mighty Sarasvatī!). Some attempts have been made to read hymn 1.133 with its *arma-ka* (=of unknown meaning, really) as a description of a ruined city (e.g. Burrow 1963, Rao 1991:32) but, in fact, the hymn mentions no ruined buildings, no fallen walls and no materials such as wood, stone or bricks! It is a ghostly scene of frightful desolation, peopled only with unfriendly she-fiends and demons (*yātumati, piśāci* and *rakṣas*). In sharp contrast the Old English poem *The Ruin* contains such persuasive details of the ancient remains (from Roman times?) that some scholars think it refers to the town of Bath (Mitchell & Robinson 1996:252-5).

e) No cotton *Karpāsa* appears in the *RV* although this plant was extensively cultivated in the ISC and the fabric was exported as far as Egypt in the middle of the 3rd millennium while the Mesopotamians adopted the name as *kapazum* (? from prākṛta *kapāsa*). The *RV* has 'skin' *eta* (1.166.10; *ajina* in *AV* 5.21.7 etc), 'wool' *avi* (*RV* 9.78.1) and *śāmulya* (10.85.29) and numerous terms for clothing and weaving but no mention of cotton. Be it noted that *karpāsa* is the only word for cotton in Sanskrit.

f) Silver *rajata* also makes no appearance in the *RV* though gold and copper are well attested and silver is plentiful in the ISC.

The word *rajata* occurs in *RV* 8.25.22 and it denotes a steed or a chariot 'shining white'. Only in later texts is it used singly (*AV* 5.28.1) or with *hiraṇya* to denote 'silver' or 'white gold' = 'silver' (see *Vedic Index* 2: 196-7 and Lubotsky 2: 1169).

g) Rice *vrīhi* too is absent from the *RV* although it appears in various sites of the ISC from at least 2300 (and in the Ganges Valley from the 6th millennium)[12]. The *RV* knows only *yava* 'barley'. Rice becomes important in post-Rigvedic ritual and the more general diet.

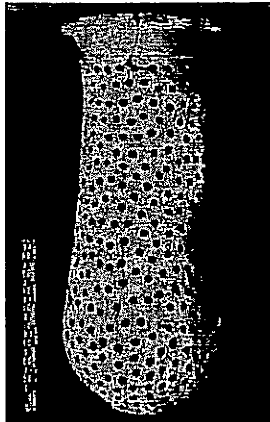
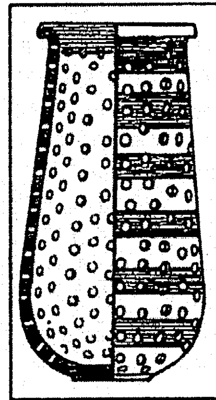
h) Literacy is not known in the *RV*. D. Frawley thought that a passage in *AV* 19.72 may refer to writing: 'From whichever receptacle *kośāt* we have taken the Veda, into that we put it down'. Books in ancient India consisted in collections of palm-leaves or strips of birch-bark and were kept in boxes (1991: 249). Whether this is enough to establish knowledge of writing is doubtful. But in any event, writing was known in the ISC but not in the *RV*.

i) The *RV* has no allusions to artistic iconography – paintings, relief representations, statue(-tte)s or seals, all so common in the ISC. (The *RV* 4.24.10 asks "Who will buy this my Indra" and this is thought by some to refer to a statuette, but this could be a transfer of favour and it is the only reference in the whole *RV* without the use of any word for statue or icon.)

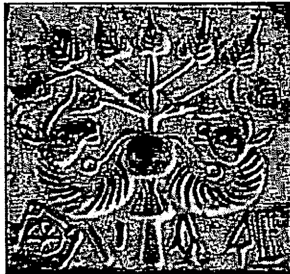
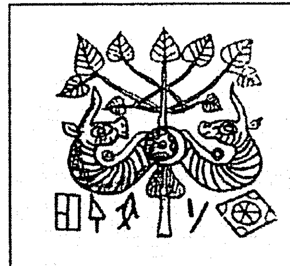
13. Now all the fore-mentioned features are found in post Rigvedic texts – the Saṃhitās, the Brāhmaṇas and fully in the Sūtra literature. For instance, brick altars are mentioned in *Śatapatha Brāhmaṇa* 7.1.1.37, or 10.2.3.1 etc. Rice *vrīhi* is found in *AV* 6.140.2; 7.1.20; etc. Cotton *karpāsa* appears first in Gautama's (1.18) and in Baudhāyana's (14.13.10) *Dharmasūtra*. The fact of the convergence of the post-Rigvedic texts and the Harappan culture was noted long ago by archaeologists. B. and R. Allchin stated unequivocally that these features are of the kind "described in detail in the later Vedic literature" (1982: 203).

If we had a case of 2 or 3 items we could bypass them saying indeed that absence of evidence is no evidence of absence; but the items are many and the absence of the building material *iṣṭakā* in the *RV* (against the presence of wood, stone, mud and metal) seems to me quite decisive.

We must also take into account that many iconographic motifs, Harappan artefacts, decorations or seals, show affinity with elements found in post-Rigvedic texts. Thus PK Agrawala (2005) draws attention to round-bottomed perforated pots from Harappan sites and a vessel (*kuṃbha)* with nine holes (*navavitṛṇṇa)* or 100 holes (*śatavitṛṇṇa)* mentioned in *Śatapatha Br* 5.5.4.27 and *śatatṛṇṇā kumbhī* 'a pitcher with 100 perforations' in 12·7·2·13. These and other similar descriptions echo the White Yajurveda (*Vājasaneyī Saṃhitā*)verse 19: 87 'a pitcher with 100 streams'. Such vessels were used for ritual sprinkling. A second parallel is furnished by the two-horned bovine-like animal, duplicated and facing itself, on a Mohenjodaro seal with long necks and the pipal tree growing out of their juncture. This corresponds (writes Agrawala) to the two-headed cattle *dvāyā* in *AV* 5.19.7. Agrawala mentions other parallels of a two-headed tiger and a two-headed bird (2005: 10 - 13). Thus it is indeed the later Vedic texts that have parallels with the Harappan arts and crafts, not the *RV*.

Pot with 100 perforations: *Śatapatha Br 12.7.2.13.*

dvāyā in *Atharvaveda 5.19.7;* symbol of OM?

N Rajaram follows a similar line of investigation (2005) and adduces iconographic motifs from Harappan seals but does not distinguish between Rigvedic (or early Vedic) and post-Rigvedic (or late Vedic) periods and cultures. Thus he cites the aśvattha-

tree foliage and the horse and links them to the *RV* disregarding the fact that these are not exclusively Rigvedic and therefore could be simply from a post-Rigvedic period. Moreover he cites the svastika and the sign of sound OM, both of which are, as he states, from post-Rigvedic texts. It is also doubtful that the seal-sign which is horizontal, depicts the symbol for OM, which should be upright, anyway. (But, of course, the seal could be seen as upright since we don't know what the symbols indicate.)

On this score, then, the bulk of the *RV*, is pre-Harappan, or 3000+. This, of course could well be the very early proto-ISC (=Harappan culture) starting at c3800 (Gupta 2007).

14. Some Brāhmaṇas comment on or give explanations of incidents in the Rigvedic hymns. For example, the *Aitareya Br* 7.13.33 narrates extensively the story of Śunaḥśepa, alluded to briefly in *RV* 1.24.12-13 and 5.2.7, while the *Śatapatha Br* 11.5.1 comments on the the Pururavas and Urvaśī love story given elliptically in dialogue form in *RV* 10.95. Another point to note. Some legends in the *RV* remain unexplained. For instance, who was Bhujyu whom the Aśvins saved from a tempest (1.116.3-4; etc) and how did he find himself in that predicament? The later texts say nothing more. Or, take the birth of Indra; was he an unwelcome child to Aditi and did he commit parricide (4.18.1ff; etc)? Again, the Brāhmaṇas tell us nothing. Obviously such exegetical texts would not be composed until the understanding regarding the older texts has lapsed: this implies many centuries.

What is the date of the Brāhmaṇas? Well, S. Kak ascribes the *Śatapatha* to the early third millennium (1997, 1994) saying that the Pleiades/Kṛttikas not swerving from the east, as is stated in this *Brāhmaṇa* (2.1.2.3), could only occur c2950 BCE. Narahari Achar (Prof. Astrophysics, Memphis, USA) using computer and planetarium apparatus arrived at a date c 3000 pointing out that S. B. Dikshit had arrived at similar conclusions 100 years earlier but was ignored by Western scholars (1999): these scholars really had no knowledge of astronomy. B. B. Lal too thought that "astronomical calculations (re *Aitareya Br*) would place the *RV* in

the fourth millennium" but did not then adopt such a date only because he could not judge the evidence (1997: 286).

So we could again place the *RV* easily before 3000.

15. The *Bṛhadāraṇyaka Upaniṣat* has three *vaṃśas*, i.e. list of teachers, each comprising 65-70 names. The first 4 or 5 are names of gods in the normal Indic way which ascribes the beginning of every human activity to some deity. Let us take the mainstream date of early sixth century for this text (say 550) according to the AIT premises and let us take 60 teachers giving to each one an average of 15 years, though 20 and 30 would be more realistic. These calculations (60×15=900 plus 550) give us a date c1450 for the inception of the doctrines in this Upaniṣat. The chief doctrine is that the self of man (*ātman*) is the same in all beings and the same as the Self of the universe (*brahman* 'Mystic Spirit' or 'Absolute'). We should also bear in mind that the teachers' names are quite different from those of the seers of the *RV* hymns as given by the native tradition. Following others, M. Witzel thinks these lists "rest on typically weak foundations" (2001, §19) but this is a typically weak subterfuge because the results of calculating the number of years prove how unrealistic the AIT chronologies are! Let us see, then.

The Upanishadic doctrine of the identity of the individual self and the universal self, in the formulations *ayam ātmā brahma* and *ahaṃ brahmāsmi* 'the Self is the Absolute' and 'I am the Absolute', should be known, then, c1500 or 1300 or 1200, when, according to the AIT, the *RV* was composed. Yet, for undetermined reasons, quite surprisingly, this doctrine is totally unknown in the *RV* in these terms (although enunciated differently) and begins hesitantly to appear in the *AV* (e.g. 10.2.32-3; 11.4.23; etc). Consequently, the *AV* should be placed at least c 1600 and the *RV* c 1800, always following the AIT assumptions. But the *RV* is composed, even according to Witzel (2005: 90), in Saptasindhu, yet the IAs do not appear in this area before, at the very earliest, 1700, and the *RV* gets composed after several centuries!

This is one of the comical paradoxes that the mainstream chronology refuses obdurately to resolve. Yet, on the grounds

given in §§12-13 above, we saw that the *RV* should be assigned to the 4th millennium, say around 3600. Now subtracting two hundred years for the *AV* hymns and the start of the Upanishadic doctrine and a further 900 or 1000 years (for the teachers) we should place the early Upaniṣads at the start of the period of the Mature Harappan, i.e. 2500 (with the *ātma-brahman* doctrine having come down orally) and the Sūtra texts immediately after. These dates satisfy yet another requirement. The word for cotton *karpāsa* is first used in the Sūtra texts as we saw in §12,e above and the cultivation of the plant (although seeds of it were found in much earlier periods) gets well established c 2500. All these dates are, of course, approximate.

16. Astrophysicist Achar pursued his **palaeoastronomical research** into the *Mahābhārata* epic also, examining astronomical references in Books 3, 5 and 18. His sky map showed that all these converge in the year 3067. (Achar 2003; see also Kazanas 2002: 295-7). Achar acknowledged that, in 1969, S. Raghavan had arrived at the same date. Now, it is obvious that the *Mbh* had acquired many accretions over many centuries and that it was streamlined stylistically perhaps first in the 2nd millennium and finally at about the start of the Common Era. It is obvious that it contains much late material like 2.28.48-9, which mentions Rome and Antiochia (*romā* and *antakhī*): this could not be earlier than 300 BCE since Antioch was founded in 301. On the other hand, the frequent use of the bow and, moreso, the use by Bhīma of a (tree trunk as a) club show much more primitive conditions.

Thus the war took place in 3067 and the core of the *MBh* in poems and songs was laid down in that year. This and the native traditional view that the Kali Yuga came at 3102 are both correct, according to Achar. He pointed out that the Kali Yuga had no full force until the death of Kṛṣṇa which occurred 35 years after 3067, at 3032 (private communication June 2006); but immorality had set in already, as is shown by the unjust behaviour of the Kauravas and some reprehensible acts during the war itself. Surely no bards (compilers or redactors) in the 3rd cent CE or the 3rd cent BCE could possibly know the star and planet positions relative to the

nakṣatras or the zodiac signs of the year 3067. The astronomical references examined by Achar (and Raghavan) are so numerous that chance coincidence has to be discarded (Achar 2003). However, that the war took place in 3137 and bards began to sing of these events two generations later should not be precluded. Personally, I still tend towards the traditional view of the War taking place in 3137. The Megasthenes report (from c300 BCE) of the ancient kings from 6000+, surviving in Arrian and other classical writers, supports these long periods of the past (Kazanas 2003). Then, deterioration in behaviour would have started in the *sandhyā* transitional period before the onset of the Kali Yuga – when also the Kṣatriyas passed away. Here a question remains: how did the astronomers (and Āryabhaṭa especially) determine the date for Kali Yuga as 3102?

17. The Sarasvatī river furnishes useful literary and archaeological evidence for dating the *RV.* It is a mighty river extolled in all Books of the *RV* except the fourth. It is *nadītamā, ambitamā, devitamā* 'best river, best mother, best goddess' (2.41.16); it is swollen and fed by three or more rivers *pinvamānā sindhubhiḥ* (6.52.6); it is endless, swift moving, roaring, most dear among her sister rivers; together with her divine aspect, it nourishes the Indoaryan tribes (6.61.8-13). In 7.95.2 the river is said to flow pure in her course "from the mountains to the ocean" *giribhyaḥ ā samudrāt.* Then 7.96.2 and 10.177 pray to the rivergoddess for sustenance and good fortune while 10.64.9 calls upon her (and Sarayu and Indus) as "great" and "nourishing". Clearly then, we have here, even in the late Bk 10, a great river flowing from the Himalayas to the ocean in the south, fed and swollen by other rivers and sustaining the tribes of the IAs on its banks – not a river known in the past or in some other region, or a river now considerably shrunk (Witzel 2001)[13].

Some scholars claim that here *samudra* does not mean 'ocean' but confluence and especially the place where a tributary flowed into the Indus. The last point can be discarded since there is not the slightest hint elsewhere that the Sarasvati flowed into the Indus – in which case the Indus and not Saraswati would have been lauded as the best river (2.41.16) We can also aver with full

certitude (as the *Vedic Index* does under *Samudra)* that the Rigvedic people knew the ocean (see Kazanas 2004a; also Prabhakar 1994). The meaning 'terminal lake(s)' adopted by Witzel is entirely fanciful. In his Dictionary M. Mayrhofer gives for *samudra* only 'confluence' and 'ocean/sea' (1996 *EWA*). And the Vedic poet would certainly have used (not *ā samudrāt* but) *ā sarobhyaḥ* 'to the terminal lakes' maintaining his – ᴗ – – cadence. This phrase would then have indicated clearly the alleged fanciful etymological connection of the name *Saras-vatī* 'she who has (terminal) lakes'. The name means rather 'she who has swirls and currents', since the primary sense of √*sṛ* (>*saras*) is 'movement' (*gatau*) or 'flowing, leaping, rushing'.

Please, consider also that the Vedic **-s-** is inherited from PIE, according to all IEnists, whereas Avestan **–h-** is a devolved, not PIE, sound. Vedic √*sṛ* has many primary and secondary cognates like *sara, sarit* etc. Now Avestan has no cognates for √*sṛ* and its products, and the Avestan noun for lake is *vairi-*; the stem *hara-* (cognate with Vedic *saras)* occurs only in the river name *Haraxvaiti.* Consequently, it is the Iranians that moved away from the Indoaryans as, indeed, is shown by their memory of having lived in a location they called *Haptahəndu* = Saptasindhu. The root *sṛ* has cognates in other IE branches, Gk *hial-*, Latin *sal-*, Toch *sal-* etc as is shown by Rix H., 1998. This, together with the fact that, like Greek, Avestan has no obvious system of roots and derivatives (see §7, above), should be enough to question if not refute various IEnists' claims that Avestan retains older forms of nouns and verbs and that therefore the Indoaryans were with the Iranians in Iran in the common Indo-Iranian period – before moving to Saptasindhu. One should also note that these linguists rely entirely on linguistic facts amenable to a reverse interpretation and ignore other aspects – literary, mythological, archaeological and genetic (for which see §19).

The river Sarasvat in Saptasindhu is thought to have dried up almost completely by 1900 (Allchins 1997: 117; Rao 1991: 77-79). In previous years it had lost tributaries to the Indus in the West and

the Ganges in the East. Is there any evidence that it flowed down to the Indian (or Arabian) ocean at any earlier period?

G. Possehl examined (1998) all the palaeoenvironmental and geological data relevant to the Sarasvat river and concluded that the river could have flowed down to the ocean only before 3200 at the very latest and, more probably, before 3800! He re-stated his finds in his study of 2002 (pp 8-9). P-H. Francfort has been just as certain of a date 3600-3800 in his survey of 1992.

All this helps us place the passages ascribing the grandeur of river Sarasvat at a date before 3200 at least. (For a detailed examination of this entire issue see Kazanas 2004a; for more recent scientific investigations through satellite showing the course of the old Sarasvat reaching the ocean see Sharma J.R. et al 2006.)

18. In a recent publication **Dr S. Levitt** (of New York), who is by no means an indigenist, examined the development of the "early Indic tradition" and the development of religion in ancient Mesopotamia.

After comparing several elements in the Vedic and Mesopotamian religions, Levitt concluded: "We can date the early Indic tradition on the basis of comparable points in ancient Mesopotamia. By this, the *Ṛgveda* would date back to the beginning of the third millennium BC, with some of the earliest hymns perhaps even dating to the end of the fourth millennium BC" (2003: 356).

However, unaware of Levitt's paper, I myself made at about that time a very detailed comparative study of Vedic and Mesopotamian religious (mythological) motifs, published in *Migration and Diffusion* vol 24, 2005[14]. In this I showed that since more than 20 motifs in the Vedic texts had close parallels in other IE branches (e.g. the horse mythology, the skyboat of the Sungod, the Flood, the elixir from heaven, the creation of cosmic parts from the dismemberment of a divine being, etc) and were therefore of Proto-Indo-European provenance, they could not have been borrowed by the Vedics from the Mesopotamians as is usually alleged (McEvilley 2002; Dalley 1998) but must be

inherited and therefore older than the Mesopotamian (Sumerian, Babylonian etc) parallels. Since the Mesopotamian culture (starting with old Sumerian) surfaces c 3000, the Vedic motifs must be earlier. Most of these have no parallels in Ugarite, Hebrew and other intermediate Near-Eastern cultures.

Thus again we arrive at a date before 3000 for the bulk of the *RV*.

19. Since, according to the preceding discussion we must now assign the (bulk of the) *RV* to c 3200 at the latest and since the *RV* by general consent was composed in Saptasindhu, then it follows that the IAs were ensconced in Saptasindhu by 3200 and that the ISC was a material manifestation of the early oral Vedic tradition expressed in the *RV*. This issue was treated by me extensively in preceding papers and no more need be said now. I should only add that, in fact, more and more scholars in the West have re-examined the issue and rejected wholly or in part the mainstream view advocating instead a movement Out of India into Europe: Schildmann 1998; Klostermaier 1998, 2000; Friedrich 2004; Hasenpflug 2006.(See also §9, above.)

To all this I should add the increasing evidence from Genetics which declares that no substantial flow of genes occurred from Europe or the northwestern adjacent areas into India before 600 BCE. On the contrary, recent genetic studies show an outflow from India into countries west and north and Europe (Sahoo et al 2006; Oppenheimer 2003). M. Danino has made a useful overview of the evidence produced by studies over the last 10 years in his paper 'Genetics and the Aryan Debate', published in *Purātattva* (2006b). Thus the Indoaryans are now shown to be indigenous to Saptasindhu and there is nothing to undemine my contention that the *RV* in its bulk is pre-Harappan. What I am arguing for primarily and most emphatically is the proposition that, since the only biological change is in the skeletal record of the area and occurs c 6000-4500 (Kazanas 2002a with references), if the Indoaryans made an entry it would have happened at that period; but by 1700 BCE they certainly were indigenous. Genetic evidence, however, suggests strongly now an Out-of-India scenario.

Cavalli-Sforza and his team state that "Indian tribal and caste populations derive largely from the same genetic heritage of Pleistocene [=10 000 to 3 mya] southern and western Asians and *have received limited gene flow from external regions since the Holocene* [=c 10 000 to present]. The phylogeography [=neighbouring branches] of the primal mtDNA and Y-chromosome founders suggest that these southern Asian Pleistocene coastal settlers from Africa would have provided the inocula for the subsequent differentiation of the distinctive eastern and western Eurasian gene pools" (emphasis and square brackets added; Cavalli-Sforza 2003).

Another geneticist, S. Oppenheimer, offers independent confirmation (2003) that there was no Aryan entry, either male or female; he focuses on the M17, or so-called "Caucasoid" (=Aryan!), genetic marker: "South Asia is logically the ultimate origin of M17 and his ancestors; and sure enough we find highest rates and greatest diversity of the M17 line in Pakistan, India and eastern Iran, and low rates in the Caucasus. M17 is not only more diverse in South Asia than in Central Asia but diversity *characterizes* its presence in isolated tribal groups in the south, thus **undermining any theory of M17 as a marker of a 'male Aryan invasion' of India**" (2003: 152: my emphasis). He adds that this M17 marker travelled from India or Pakistan (= our Saptasindhu) through Kashmir, Central Asia, Russia and then Europe after 40 000 BP (p 154). Thus migration is from east westward.

A more recent study (Sahoo ... Endicot, Kivisild ... Kashyap 2006) concludes: "The Y-chromosomal data consistently suggest a largely South Asian origin of Indian caste communities and therefore argue against any major influx, from regions north and west of India" (p 843); then again: "It is not necessary, based on the current evidence, to look beyond South Asia for the origins of the paternal heritage of the majority of Indians at the time of the onset of settled agriculture. ... our findings do support a local origin of haplogroups F* and H" (p 847).

However, all these finds are by no means final and we need to be cautious; but the indications now are that there will be no

reversal. If there was a migration, this was most probably out of Saptasindhu and the adjacent region.

Why mainstreamers insist on the AIT is a mystery. Lord C. Renfrew wrote of the AIT (1989:182): "this comes rather from a *historical assumption* about the 'coming' of the Indo-Europeans" (my emphasis). Then Edmund Leach wrote (1990): " Because of their commitment to a unilateral segmented history of language development that needed to be mapped onto the ground, the philologists took it for granted that proto-Indo-Iranian was a language that has originated outside India or Iran... From this we derived the myth of the "Aryan invasion"." These are the two legs of clay upon which stands the AIT and its variants.

Leach went further saying that after the discovery of the Indus-Sarasvati Civilization "Indo-European scholars should have scrapped all their historical reconstructions and started again from scratch. But this is not what happened. Vested interests and academic posts were involved" (1990). This is still true. But the new genetic evidence will soon perhaps force linguists to reconsider their theories. And we must not forget that there may well have been an IE continuum from the Steppe to Saptasidhu and the IAs did not move from their location. It is worth noting that S. Zimmer admitted (2002) that (although himself a mainstream non-indigenist) he could not be certain of the exact location of the PIE homeland since the facts are so obscure in those far-off times. More recently, H-P Francfort, the eminent excavator of Shortughai, expert on Central Asia Oxus area (or BMAC) and N-W India, critiqued V. Sarianidi, E. Kuzmina and J. Mallory and their theories about [proto-]Indo-Iranian movements through Oxus region (2005: 262-8); further on (p 283 ff) he pointed out that the pantheon in the Oxus iconography has a dominant goddess and so does not tally with Iranian and Indoaryan religions: on the whole he is most reluctant to accept Indo-Iranians (or Aryans) passing through that area c 1800-1400 BCE. So even some mainstreamers have now serious doubts about the alleged Aryan immigration/invasion.

20. Does it matter whether the *RV* was composed in the 4th millennium or c 1000? whether the Indoaryans are indigenous or came to India c 1500? After all life goes on irrespective of such ancient events. So from one point of view, no, it doesn't matter. But from another, it does matter. First, it is the function of scholarship to establish and promote true knowledge so that our life be regulated by this - not prejudices, partisan views (even patriotic but false) or pet theories. Second, Indian (proto-)history must be restored and revalued in a correct time-frame. It is sad, indeed, as Frawley wrote (2002 *The Hindu* June 18th), "to note how intellectuals in India are quick to denigrate the extent and antiquity of their history". Third, the *RV* contains ideas that are of great value to mankind and reveal, as other recent discoveries in prehistoric cultures (Rudgley 1998 passim), that so-called "primitive" people had much important knowledge that held them at a fairly high level of civilization enabling them to live in harmony with the natural processes (*ṛta*) of their environment. The *RV*, e.g., preserves for us, from that so ancient period, the idea of a Primal Unity that is the First Cause of the universe and all its phenomena: not only in the *Nāsadīya sūkta* and the *Puruṣa sūkta* (10.129 and 90) but also in 1.164,6; 3.55 refrain; 6.75.19; 8.58; 10.114.5 and less obviously in others. This concept is absent from all other IE traditions (Hittite, Greek, Roman etc) and may well have been an essential constituent of PIE culture lost in the other branches. Thus the Rigvedic IA culture (and perhaps PIE) consisted not so much of material artefacts but of inner knowledge and spiritual strength - *bráhma várma mamā́ntaram* 'the holy-power is my inner defence' *RV* 6.75.19.

Let us hope that the noxious AIT and all notions rooted in it will sooner rather than later end up in the only place they should be - the dustbin of History. Let us also hope that the *RV* will be re-examined not only for its historical information but its cultural ideas that will illuminate many other studies and disciplines in Ancient History, Anthropology, Civilization, Linguistics, Religion, Sociology, etc.

Endnotes

[1] IE = IndoEuropean; PIE = ProtoIndoEuropean; IA = Indoaryan(s). A large section of this paper is a revised version of a talk given at a Conference at the Center of Indic Studies, University of Massachusets, in June 2006.

[2] All dates are BCE except where stated otherwise. Dates in brackets, e.g. (2001) denote a publication by a modern author and will be found in the Bibliography.

I use the name Saptasindhu simply as a *bahuvrīhi* adjective. The 7 rivers are frequently mentioned in the *RV* - 1.32.12; 34.8; 35.8; etc, etc. In the *Avesta* it is said that the Old Iranians had, before settling in Iran, been to various places and one of them was *Haptahəndu*. I take it that *Haptahəndu* indicates Saptasindhu, a name which has been frequently used by others before me.

[3] Take as example the hospital. Here the patients are far more numerous than the doctors. Numerically they are more significant but it is the doctors who do the truly significant work, restoring health.

[4] A deep and wide inconsistency between thought and action must indicate something is amiss.

[5] She points out that, e.g., there are 36 sound-rules governing consonantal differences between Latin and Germanic but only 34 attested parallels to support them. Marcantonio is, however, much too rigid in laying down her four criteria for scientifically determining genuine correspondences. Her 2nd criterion - that "a given sound in language A can correspond with only one sound in language B, or perhaps two (similar) alternatives in language B" contains an assumption which is untrue, the assumption that languages change uniformly. Changes in different languages are neither uniform nor parallel, as she herself points out (Marcantonio & Brady 2005), and therefore one can't rely on such strict correspondences. Ultimately one must rely on "naked eye" as she puts it, or "impressionistic" perceptions. If some regular sound-correspondences can be established, all well and good; but they should not be taken as absolute and exclusive of other alternatives. Thus Gk *theos* 'god' could well = L *deus*, S *deva*, and S *dvār* 'door' could well = Gk *thura*, Gmc *dawr/tor* etc. Marcantonio cites also S *napāt* 'offspring', Gk *nepodes* 'children, offspring', L *nepōs* 'nephew' and Gmc *nefo* 'nephew, relative' and takes exception to the sound *d* in Gk *nepodes* (pl with unattested sing) observing that it should - according to the specific "law"- be *t*. Well, as I said, the changes were not uniform and one must rely on

one's discrimination, not "laws" nor statistics, though both are often useful.

6 Here too we have examples showing apparently great confusion as noted in §7.

7 In a private communication, Nov 2004, Mallory mentioned the Jews who wandered much and long and preserved their traditions of the *Old Testament*. If we take the Hebrew traditional orthodox view (which is by no means accepted generally), the Judaic people had literacy certainly since the time of Moses (c 1300-1100?) and most probably since their stay in Ur of the Chaldees c 1700-1600 (*Genesis* 11), since the Mesopotamians had literacy for more than a millennium earlier. So we can't say that the Jews maintained an oral tradition during their long travels. Then, the Jews mention constantly the (mis-)adventures they had and the different people they met on their travels. If we assume that the IAs had their oral tradition even as (according to the AIT) they travelled from the Steppe to Saptasindhu, what was it they were transmitting? The Vedic texts were composed (it is universally agreed) in India. So what were they transmitting before?... Why is there not one mention in the Vedic texts of dangers, mishaps and alien people met on the way – as we find in the Judaic texts? There is no such mention for the simple reason that, unlike the Jews, the IAs did not migrate.

8 S. Talageri should perhaps be included but despite having some very good ideas, this author knows no Sanskrit and has no training in Archaeology or other related disciplines and so goes astray constantly.

9 A common Indo-Iranian period there definitely was but this, according to a different view, is to be found in Saptasindhu whence later the Iranians left.

10 Witzel's earlier views on these loans and Bryant's emphasis on linguistic argument (2001) have been contested by me in papers 2002b and 2003.

11 Bisht, for example, cites the numerous occurrences of *pur* as evidence that the *RV* knew of towns and fortifications (p 410ff). So does D. Frawley (2002: 13). This is wrong, of course: *pur* in the *RV* never means 'fort' or 'town' but a magical, esoteric defence. Merely citing verses in the *RV* but not analysing and exploring their meaning is not very sound scholarship. Here are some examples merely cited by both writers. In II, 35 the Offsring-of-Waters *Apā́ṃ Nápāt* is protected in *āmā́su pūrṣú* (usually translated as 'forts of unbaked bricks' even though bricks are nowhere mentioned!); the phrase means 'non-artificial natural defences' (in this case obviously 'divine'). In I 166, 8 Agastya prays to the Maruts to protect (*rakṣ-*) from evil *agha* and injury *abhihruti* the man they favour (*av-*) with 'hundredfold defences' *śatábhujibhiḥ pūrbhíḥ* : no such many-walled structure existed (allowing for the hyperbole) and, moreover, it is difficult to see how people can be

protected from evil by 100fold fort/towns! VI 48, 8 is a similar prayer to Agni to protect (*pā-*) from anxiety *aṃhas* the fire-kindler with '100 purs': clearly one cannot be protected from anxiety with 100 forts/cities (again allowing for a hyperbole) but only with internal occult defences. A final example: X 101, 8 is a prayer to Viśvedevās "to fashion inviolable metal purs" to protect the rite of soma-pouring; obviously no forts or towns are meant to be built around the sacrificial locus but magical other-wordly defences against demons. Not one *pur* is built and not one is destroyed by human beings! See Kazanas 2003: 224-5 and n 12; also <http://www.omilosmeleton.gr/english/en_index.html> '*Ṛgvedic pur*'; also Kazanas 2006.

12 Private communication from S.P. Gupta, Chairman of Indian Archaeological Society (June 2006). See also Sharma 1980 for rice in the Ganges basin in the 5th millennium and R. Tewari et al in *Purātattva* 2006 (vol 36: 68-75) for rice, again in the Ganga basin, north-east, in the district Sant Kabir Nagar (UP) in 7th millennium. These locations are too far from Vedic Saptasindhu but it seemss likely that by 3000 the composers of the *AV* had become acquainted with rice *vrīhi* and later this grain was cultivated in the ISC too.

13 The mainstream view (Witzel's really) that the Vedic river is merely a memory of the Iranian Haraxvaiti which belongs to the common Indo-Iranian period, when the Iranoaryans lived together in Iran before the IAs moved further southeast (according to the AIT), is no more than modern myth-making. Mainstreamers often invoke Occam's razor (i.e. that the simpler solution is more probable) but here they forget it and prefer their own complex scenario.

14 This was badly printed and the Sanskrit transliterations are unreadable! A revised version was published in *Adyar Library Bulletin* 2007.

References

Achar, B.N.N.

2003 'Date of the Mahābhārata War...' in *Quarterly Journal of the Mythic Society*, vol XCIV and Issue Nos 1-2.

2001 'Planetarium software and the date of the Mahābhārata war' paper read in Montreal, sent to N. Kazanas privately.

1999 'On exploring the Vedic sky ...' *Electronic Journal of Vedic studies* 5-2 (3-9).

Agarwal, V.

2006 'Is there evidence for the Indo-Aryan Immigration to India?' talk at WAVES Conference August. See <http:/ www.omilosmeleton.gr/english en_index.html>.

Agrawal, A.

2005 *In search of Vedic Harappan Relationship* N. Delhi, Aryan Books Intern.

Agrawal, D.P.

2002 'Comments on "Indigenous IndoAryans"' in *Journal of IE Studies* vol 30 (129-135).

Agrawala, P.K.

2005 'Two Harappan Motifs & their Vedic Affinity' in A. Agrawal ed. (10-13).

Allchin, B. & R.

1997 *Origins of Civilization* Delhi, Viking Penguins.

1982 *The Rise of Civilization in India & Pakistan* Cambridge, CUP.

Allchin, F. & Joshi J.

1995 (Eds) *Excavations at Malvan* Delhi, Archaeological Survey of India (Memoir No 92).

Andersen, E.R.

1999 'Horse-sacrifice and Kingship...' *JIES* vol27, 379-393.

Aurobindo, Shri

1914-1920 *On the Veda* and *The Secret of the Veda* articles in *Arya*, India (1956 reprinted in book form, Pondicherry).

Beveridge, W.

1968 *The Art of Scientific Investigation* (1954) London, Heinemann.

Bisht, R.S.

1999 'Harappan and *Ṛgveda*: Points of Convergence' in G. Pande, ed 1999 (393-442).

Bryant, E.

2002 'Somewhere in Asia and no more' *JIES* 30 (136-48)

2001 *The Quest of the Origins of Vedic Culture* Oxford, OUP.

Burkert, W.

1992 *The Orientalizing Revolution* Camb (MA)/London, HUP.

Burrow, T.

1973 *The Sanskrit Language* (1955 rev ed) London, Faber.

1963 'On the significance of the terms *arma-*, *armaka-*', *Journal of Indian History* vol 41 (159-68).

Cardona, G., Hoeningswald H.& Senn A.

1970 *Indoeuropean and Indoeuropeans*, Philadelphia, Univ. Pensylvania Press.

Cavalli-Sforza et al

2003 'The Genetic Heritage of the Earliest Settlers ...' *American Journal of Human Genetics* 72 (313-32).

Chakrabarti, D.K.

1999 *India: An archaeological History* Cambridge, CUP.

Chatterji, S.K.

1968 *Balts and Aryans* Simla, Indian Institute of Advanced Study.

Clackson, J.

2007 *Indo-European Linguistics: an Introduction* Cambridge, CUP.

Cohen, I.

1985(2001) *Revolution in Science* Cambridge Mass, HUP.

Dales, G.

1966 'The Decline of the Harappans' *Scientific American* 241 (92-99).

Dalley, S.

1998 ed *The Legacy of Mesopotamia* Oxford, OUP.

Dange, S.A.

1967 'The Gaviṣṭi (Go-iṣṭi)...' *Nagpur University Journal* (India).

Danino, M.

2006a *L'Inde et l 'invasion de nulle part...* Paris, Les Belles Lettres

2006b 'Genetics and the Aryan Debate' in *Purātattva*, vol 36.

2006c 'The Horse and the Aryan Debate' *Journal of Indian History and Culture* vol 13(33-59), Chennai, India.

Darmesteter, J.

1965 *The Zend-Avesta* (1883) Sacred Books of the East 4, 23, Delhi, M. Banarsidass.

Das, R.P.
2002 'Indra und Śiva/Rudra...' in Das & Meiser (eds) Geregeltes Ungestüm *Veröffentlichungen zur Indogermanistik und Anthropologie*, Bremen, Hempen Verlag.
1995 'Hunt for foreign words in *RV*' *Indo-Iranian Journal* 38 (207-238).
Day, J.V.
2001 *Indo-European Origins* Washington, Institute for the study of Man.
1994 'The concept of the Aryan race in the 19th-century scholarship' *Orpheus* 4, 15-48.
Dhavalikar, M.K.
1995 *Cultural Imperialism...* Delhi, Books & Books.
Dhar, L.
1930 *The Home of the Aryas* Delhi, Delhi Univ. Press.
Dixon, R.
1997 *The Rise and fall of languages* Cambridge, CUP.
Dumézil, G.
1954 *Rituels indo-européens* Collection 'Etudes et Commentaires', vol19, Paris, Klincksieck.
1952 *Les Dieux des Indo-Européens* 'Mythes et Religions'vol 29, Paris, PuF.
Dunstan, W.E.
1998 *The Ancient Near East* Orlando (FA), Harcourt Brace.
Edda
1996 S. Sturluson in Everyman series (1987), London, J Dent.
Elst, K.
1999 *Update on the Aryan Invasion Debate* Delhi, Aditya Prakashan.
1993 *Indigenous Indians* Delhi, Voice of India.
Erdosy, G.
ed 1995 *The Indo-Aryans of Ancient South Asia*, Berlin/NY, Guyter.
Estell, M.
1999 'Orpeus and Ṛbhu revisited' *JIES* 27 (327-333).
1985 'Primary and secondary homeland...' in *JIES* pts 1-2 (185-202).
Feuerstein, Frawley, D. Kak, S.
1995 *In search of the Cradle of Civilization* Wheaton/Madras Quest Books.

Fortson, B.
2004 *Indo-European Language and Culture* Oxford, Blackwell.
Fox, A.
1995 *Linguistic Reconstruction. An Introduction...* Oxford OUP.
Francfort, P-H.
2005 'La Civilisation de l' Oxus et les Indo-Iraniens et Indo-Aryens en Asie Centrale' in Fussman G., Kellens J., et al (below).
1992 'Evidence for Harappan irrigation system' in *Eastern Anthropologist* vol 45 (87-103).
Frawley, D.
2002 'Witzel's vanishing ocean...' http://voiceofdharma.com/indology/ReplytoWitzel.html
1994 *The Myth of the Aryan Invasion* Delhi, Voice of India (repr 2000).
1991 *Gods Sages & Kings* Salt Lake City, Passage Press.
Frawley, D. & Rajaram N.
1997 *Vedic Aryans & Origins of Civilization* Delhi, Voice of India.
Fraser, A.
1995 *The Gypsies* Oxford, Blackwell.
Friedrich, H.
2004 'A Diffusionist's view on the concept of "Language Families"' *Migration & Diffusion* vol 5, No 19 July-September (6-14).
2003 'Hebrew Ethnogenesis and Diffusion' *Migration & Diffusion* 4, 16 (Oct-Dec).
Fussman, G., Kellens J., Francfort H-P., Tremblay X.
2005 *Aryas, Aryens et Iraniens en Asie Centrale* Paris, College de France, Publications de l' Institut de Civilisation Indienne.
Gimbutas, M.
1985 'Primary and Secondary Homeland of the Indo-Europeans' in *JIES* (185-202).
1970 'Proto-Indo-European Culture' in G. Cardona & H. Hoeningswald & A. Senn eds *Indo-European & Indo-europeans* Philadelphia, Univ Pennsylvania Press.
Gnoli, G.
1980 *Zoroaster's Time and Homeland* Naples, Instituto Universitario Orientale (Series & Minores VII).
Gray & Atkinson, Q.
2003 'Language Tree Divergence Time...' *Nature 426* (435-9).

Gupta, S.P.

2007 'City in Protohistoric India' in G. Pande ed *Golden Chain of Civilizations...* PHISPC vol 1, pt 4, (pp 209-355), Delhi, Centre for Studies in Civilization.

2005 'The Saraswatī and the Homeland of Early Rigvedic Ṛishis' in A. Agrawal ed (55-73).

Hale, M.

2004 'Avestan' in R.D. Woodard (ed) *The Cambridge Encyclopaedia of the World's Ancient Languages* (pp 742-763).

Held, W., Schmalstieg W., Gertz J.

1987 *Beginning Hittite* Columbus (Ohio), Slavica.

Heskel, D.L.

1984 'Iran-Indus Valley connections' in B.B. Lal & S.P. Gupta (p 333-346).

Hock, H.H.

1999 'Out of India? The Linguistic Evidence' in J. Bronkhorst & M. Deshpande (eds) *Aryan and Non-Aryan in South Asia*, Camb Mass, HOS (1-18).

1996 'Pre-ṛgvedic convergence' in Houben E. M. *Ideology and Status of Sanskrit* Leiden, Brill.

1991 *Principles of Historical Linguistics* (2nd ed) Berlin/NY, de Gruyter.

Hopkins, E.W.

1898 'The Punjab and the Rig-Veda' in Journal of the *American Oriental Society* vol 19 (19-28).

Hughes, R. & Row, M.

1982 *The Colouring, Bronzing and Patination of Metals* Oxford, Arts Council.

Huld, M.E.

2002 'Linguistic Science, Truth and the Indocentric Hypothesis' *Journal of IE Studies* 30 (149-60).

Jarrige, J.F. & Meadow, R.

1980 'Antecedents of Civilization in the Indus Valley' in *Scientific American* August issue.

Kak, S.

1997 'Archaeoastronomy & Literature' October issue of *Current Science.*

1994 *The Astronomical Code of the Rigveda* (revised ed 2000) Delhi, Manoharlal.

Kazanas, N.

2007a 'Vedic, Mesopotamian and Egyptian Religiophilosophical Thought' in *Golden Chain of Civilizations, PHISPC*, Delhi & Simla.

2007b 'Coherence & Preservation in Sanskrit' *Research Bulletin* of the Vishveshvarand Vedic Research Institute, vol 5.

2006a 'Diffusion of IE Theonyms' in *Quarterly Journal of the Mythic Society* (Bangalore, India) vol 97, no 1 (1-29).

2006b 'Ṛgvedic *pur*' in *Man and Environment* vol XXXI (1) pp121-125.

2005a 'Vedic and Mesopotamian Cross-influences' in *Migration & Diffusion* (Vienna) vol 6, no 24.

2005b 'The Ṛgveda and Harappa' in *Vedic-Harappan Relationship* ed A. Agrawal, Delhi, Voice of India.

2004a 'Samudra & Sarasvatī in the *RV*' in *QJMS* vol 95 (90-104).

2004b 'Sanskrit & Proto-Indo-European' *Indian Linguistics* vol 65 (75-100).

2004c 'Homer, Hesiod and the Mahābhārata' in *Vishveshvarand Vedic Research Institute Research Bulletin* vol 4.

2003 'Final Reply' *Journal of IE Studies* vol 31, 1&2 (187-240).

2002a 'The 3 Functions and Unity in the *Rigveda*', JICPR 29-1 (75-98).

2002b 'Indigenous Indoaryans & the Ṛgveda' *Journal of IE Studies* vol 30 (275-334).

2001a 'A New Date for the *Ṛgveda*' in Pande G. & Krishna D. (eds) *Journal of Indian Council for Philosophical Research* Special Issue.

2001b 'Indo-European deities and the Ṛgveda' *Journal of Indo-European Studies* 29 (257-93).

1999 'The *Ṛgveda* and Indoeuropeans' ABORI 80 (15-42).

Kenoyer, J.M.

1998 *Ancient Cities of the Indus Valley Civilization* Karashi Oxford, OUP.

Klostermaier, K.

2000 *Hinduism* Oxford, Oneworld Books.

1998 'Questioning the Aryan Invasion Theory ...' *Iskcon Communications Journal* 6:1 (5-16). Republished in *International Quarterly of Indian Philosophy* June 2001 (63-78).

Kochhar, R.

2000 *The Vedic People* Hyderabad, Orient Longman.

Krishnamurti, B.

2003 *The Dravidian Languages* (Cambridge Language Surveys) Cambridge, CUP.

Kuiper, F.B.J.
1991 *Aryans in the Ṛveda* Amsterdam Atlanta, Rodopi.
Kuhn, Th.
1970 *The Structure of Scientific Revolutions* Chicago, Chicago Univ Press.
Kurylowicz, J.
1956 *L'apophonie en indo-européen* Breslau.
Kuzmina, E.E.
2002 'Comments on Kazanas' *Journal of IE Studies* 30 (161-69).
Lal, B.B.
2005 *The Homeland of the Aryans* N. Delhi, Aryan Books International.
2002 *The Sarasvatī Flows On*, Delhi, Books International.
1998 *India 1947-1997: New Light on the Indus Civilization* Delhi, Aryan Books Intern.
1997 *The Earliest Civilization* Delhi, Aryan Books.
Lal, B.B. & Gupta, S.P.
eds 1984 *Frontiers of the Indus Civilization* Delhi, Indian Archaeological Society.
Leach, E.
1990 'Aryan Invasions over four millennia' in E. Ohnuki-Tierney (ed). *Culture Through Time*, Stanford, Stanford University Press (227-45).
Lehmann, W.P.
1993 *Theoretical Bases of Indo-European Linguistics* New York, Routledge.
Levin, S.
1991 'Full and other key words shared by Indo-European and Semitic' in Lamb SM and E.D. Mitchell (eds) *Sprung from Some Common Source* Stanford (CA) Stanford University Press (166-177).
Levitt, S.H.
2003 'The Dating of the Indian Tradition' *Anthropos* vol 98 (341-59).
Lockwood, W.B.
1969 *Indo-European Philology* London, Hutchinson University Press.
Lubotsky, A.
1997 *A ṚgvedicWord Concordance* New Haven, American Oriental Society.

Mac Canna, P.
1983 *Celtic Mythology* London, Chancellor Press.
MacDonell, A.
1898 *Vedic Mythology* (repr 1995) Delhi, M. Banarsidass.
Mallory, J.
2002 'Indigenous Indo-Aryans: The Preservation and Total Distribution Principles', *JIES* vol 30 (170-182).
1997 'The Homeland of the Indo-Europeans' in R. Bleach and M. Sprigg (eds) *Archaeology & Language* London/NY, Routledge (vol 1: 93-121).
1989 *In Search of the Indo-Europeans* London, Thames & Hudson.
1973 'A Short History of the Indo-European Problem' *JIES* vol1: 21-65.
Marcantonio, A.
2002/05 'Linguistic Palaeontology: science or fiction?...in B. Winsa (ed) *Finno-Ugric Peoples in the Nordic Countries* Proceedings of the 5th Roots Conference, Stockholm, Academia Tornedallensis (77-94).
2002 *The Uralic Language Family: Facts, myths and statistics'*, Oxford/ Boston, Blackwell.
Marcantonio A., Brady, R.M.
2005 'Evidence that Indo-European reconstructions are artefacts of the Linguistic method of analysis', paper delivered in Cambridge and sent to NK privately in Oct, 05.
2002/2005 'Evidence the missing concept in comparative studies' in M. Fernandez Vest(ed) *Les Langues Ouraliennes Aujourd' huit* Bibliothèque de l' École des Hautes Études, Paris Librairie Honoré Champion (117-132).
Mayrhofer, M.
1996 *Etymologisches Wörterbuches des Altindischen* Heidelberg, Carl Winter.
1956 *KEWA (Kurtzgefasstes...)* Heidelberg, Carl Winter.
McEvilley, Th.
2002 *The Shape of Ancient Thought* New York, Allworth Press.
McIntosh, J.R.
2001 *A Peaceful Realm* Boulder (Colorado), Westview.
Misra, S.S.
1992 *The Aryan Problem: A Linguistic Approach* Delhi, M. Manoharlal.
1968 *A Comparative Grammar of Sanskrit*, Greek & Hittite Calcutta.

Morgan Kelley, E.
1992 *The Metaphorical Basis of Language*. New York, Lewiston.
Mitchel, B. & Robinson, F.
1996 *A Guide to Old English* Oxford, Blackwell.
Nichols, J.
1997-98 'The Epicentre of the IE linguistic spread' and 'The Eurasian spread zone ...' in R. Blench & M. Spriggs (eds) *Archaeolology & Language*, NY/London Routledge, vol 1, 122-148 & 2, 220-266.
Onians, R.B.
1989 *The Origins of European Thought ...* Cambridge, CUP.
Oppenheim, A.L.
1977 *Ancient Mesopotamia* Chicago/London, Chicago Univ Press.
Oppenheimer, S.
2003 *The Real Eve: Modern Man's Journey out of Africa* NY, Carroll & Graf.
2002 *Out of Eden: the Peopling of the World* London, Constable.
Pande, G.C.
(ed)2007 *Golden Chain of Civilizations...* PHISPC vol 1, pt 4 N. Delhi, Centre for Studies in Civilization.
1999 *The Dawn of Indian Civilization up to 600 BC* N Delhi, PHISPC, vol 1, pt1, Centre for Studies in Civilization.
Parpola, A. et al
1969 *Decipherment of Proto-Dravidian Inscriptions...* Copenhagen, Scandinavian Institute of Asian Studies.
1988 *The Coming of the Aryans to Iran and India* Helsinki, Studia Orientalia.
Pert, C.B.
1997 *Molecules of Emotion* NY/London, Constable.
Piggott, S.
1992 *Wagon, Chariot and Carriage* London, Thames & Hudson.
1983 *The Earliest Wheeled Transport* London Thames & Hudson.
Possehl, G.L.
2002/3 *The Indus Civilization* Roman & Littlefield (Non-NBN); N. Delhi, Vistaar Publications.
1998 'Did the Sarasvati ever flow to the Sea?' G.S. Phillips et al (eds). *Arabia & its Neighbours* (in *Honour of Beatrice de Cardi*) Turnhout, Brepols (339-54).

Possehl, G.L. & Gullapalli, P.
1999 'The early iron age in South Asia' in V. Pigott (ed) *The Archaeometallurgy of the Asian old World* Philadelphia, The University Museum (153-175).
Prabhakar, C.L.
1994 'The Ancients' Approach to the Ocean' *Quarterly Journal of the Mythic Society* vol 85, No 2 (30-45).
Puhvel, J.
1970(ed) *Myth and Law among Indo-Europeans*, Berkeley, Univ California Press.
1987 *Comparative Mythology* Baltimore, J. Hopkins Univ. Press.
Rajaram, N.
2005 'Connections between Harappan Seals and the Vedic Literature' in Agrawal A. 2005 (135-41).
Rao, L.S.
2006 'The Harappan spokedwheels...' in *Purātattva* vol 36 (59-67).
Rao, R.
1880 'Puzzles for the Philologists' *Theosophist* 1 (305-308).
Rao, S.R.
1991 *'Dawn & Devolution of the Indus Civilization* Delhi, Aditya Prakashan.
1973 *Lothal & the Indus Civilization* Bombay, Asia Publications.
Renfrew, C.
1989 *Archaeology & Language* Hammondsworth, Penguins.
Rix, H.
1998 *Lexikon der indogermanischen verben...* Wiesbaden, Reichert.
Roux, G.
1992 *Ancient Iraq* London, Penguins.
Rudgley, R.
1998 *Lost Civilizations of the Stone Age* London, Century.
Sahoo, S., Singh, A., Kivisild, T... et al
2006 'A Prehistory of Indian Y-chromosomes' in *Proceedings of the National Academy of Sciences* (USA) vol 103, no 4, 843-8.
Sankalia, H.D.
1974 *Prehistory & Protohistory of India & Pakistan*, Poona, Deccan College.

Schildmann, K.
1999 'Entzifferung der Burrows Cave Texte in USA und der Glozel Texte in Frankreich als Indus-Sanskrit Texte' in *Studia Orientalia und Atlantica*, Bonn.
1998 *Indus/Burrows Cave USA-Deciphering* Bonn, Studia Orientalia et Indo-Atlantica, Fascicle 1.

Sergent, B.
1997 *La Genèse de l' Inde* Paris, Payot & Rivages.

Sethna, D.K.
1992 *The Problem of Aryan Origins* 2nd Ed, Delhi, Aditya Prakashan.

Shaffer, J.G.
1984 'The Indo-Aryan Invasions: Cultural Myth and Archaeological Reality' in Lukaks J.R. (ed) *The People of South Asia: the Biological Anthropology, of India, Pakistan & Nepal*, New York, Phenum.

Shaffer, J.G. & Lichtenstein, D.A.
1999 'Migration, philology & South Asian archaeology' in J. Bronkhorst & M. Deshpande (eds) *Aryans and Non-Aryan in South Asia* Camb Mass, HOS.

Sharma, G.R.
1980 *History to Protohistory: Archaeology of the Vindhyas and the Ganga Valley*, Allahabad, Allahabad University.

Sharma, J.R., Gupta A.K., Bhadra, B.K.
2006 'Course of Vedic river Saraswati...' in *Purātattva* vol 36 (187-195).

Sharma, R.S.
1999 *Advent of the Aryans in India* New Delhi, Orient Longman.
1996 *Looking for the Aryans* Hyderabad, Orient Longman.

Sihler, A.L.
1995 A New Comparative Grammar of Latin & Greek Oxford, OUP.

Singh, Bh.
1995 *The Vedic Harappans* Delhi, Aditya Prakashan.

Speirs, A.G.
1978 *The Proto-Indo-European Labio-velars* Amsterdam, Hakkert.

Stone, A.
1997 *Ymir's Flesh* Loughborough (Britain), Heart of Albion Press.

Swadesh, M.
1955 'Towards greater accuracy in lexico statistical dating', *International Journal of American Linguists* vol 21 (121-137).
1953 'Archaeological and Linguistic chronology...' *American Anthropologist* vol 55 (349-352).

Szemerényi, O.
1972 'Comparative Linguistics' in *Current Trends in Linguistics*, vol 9 (119-198).
1996 *Introduction to Indo-European Linguistics* (transl from German 1990, with additional notes and references) Oxford, OUP.

Talageri, S.
2000 *The Rigveda* New Delhi, Aditya Prakashan.

Tewari, R., Srivastava, R.K., Singh, K.K., Saraswat, K.S.
2006 'Further excavations at Lahuradeva, District Sant Kabir Nagar...' *Purātattva* vol 36 (68-75).

Thapar, B. & Mungal, R.
1994 'The Indus Valley (3000-1500 BC) in A Dani & J.P. Mohen (eds) *History of Humanity* vol 2, London, Unesco/Routledge (246-65).

Thapar, R.
1999 *The Aryan Question Revisited* <http://members.tripod.com/ascjnu/aryan.html>.

Thomas, P. et al
1995 'Harappan subsistence patterns...' *Man & Environment* (India) 20, 2 (33-41).

Trautmann, T.
1997 *Aryans and British India* N. Delhi, Vistaar Publications.

Tremblay, X.
2005 'Grammaire comparé et grammaire historique' in Fussman G., Kellens J., et al.

Trubezkoy, N.S.
1939 Gedanken über das Indogermanen problem in *Acta Linguistica* 1 (81-89).

Vedic Index
1995 (1912) by MacDonnell A. & Keith A., Delhi, M. Banarsidass.

Wadler, A.
1936 *Germanische Urzeit*, Basel (reprinted Wiesbaden 1980).
1988 *Der Turm von Babel* Wiesbaden, O Horrassowitz.

Watkins, C.
2001 *How to kill a Dragon: Aspects of IE Poetics* Oxford, OUP (1995).
2000 *The American Heritage Dictionary...* Boston/NY, Mifflin & Co.

Winternitz, M.
1981 *History of Indian Literature* (transl by V. S. Sarma) 3 vols, Delhi, M. Banarsidass.

Witzel, M.
2005 'Central Asian Roots and Acculturation in South Asia' in Osada Tosiki ed *Linguistics, Archaeology and the Human Past,* Kyoto, Indus Project, Research Institute for Humanity and Nature.
2003 'Ein Fremdling in Ṛgveda' *JIES* vol 31.
2001 'Autochthonous Aryans' *Electronic Journal of Vedic Studies* 7-3, 1-93.
1995 'Early Indian History' and 'Rigvedic History' in G. Erdosy (ed) *The Indo-Aryan of Ancient South Asia* (85-125) Berlin/NY, De Guyter.

Zimmer, S.
2000 'The easy way to truth...' *JIES* 30, 192-201.

Evidence that most Indo-European Lexical Reconstructions are Artefacts of the Linguistic Method of Analysis

Angela Marcantonio
University of Rome 'La Sapienza'

A typical book on Indo-European linguistics takes for granted that the evidence for the Indo-European family tree is 'obvious' or 'compelling'. In more specialist research books, one finds a focus on how to 'fit' the data to the model. The earliest example of this was Verner, who famously described 'exceptions' to Grimm's Law. Verner modified the model by introducing 'contextual specifications', thus converting the exceptions into 'apparent exceptions'. Over the years, the counter-evidence to the model has been 'explained away' with similar processes, so that today one can find very little direct counter-evidence to the model. Here I examine this process critically. I ask if the linguistic method has become circular – in other words, I ask: "Have we fitted the data to the model, or have we made the model so flexible that it can fit almost any data?"

1. Introduction

1. 1. The circularity issue

The objective of this chapter is to address the general issue of the flexibility, and related risk of circularity, which is embedded into the comparative method with the purpose of verifying whether, and to which extent, these weaknesses may have a negative impact on the results it yields, in this case the IE comparative corpus (and, arguably, the IE theory as a whole). Following the results of this investigation, I shall argue that the great majority of the conventionally stated IE sound laws lack statistical significance

and that, as a consequence, most of the conventionally established correspondences (within a chosen corpus) are not correspondences, but similarities, most probably 'chance resemblances'.

Among the weaknesses[1] embedded in the comparative method, that of circularity has since long been identified, and can be defined as follows (see also Fox[2] [1995]). 'Circularity' is where we:

a) assume that a set of words are of IE origin
b) make a reconstruction of the origin and history of the words in question through the (assumed) sound-changes associated with this history (on the assumption that the words are of IE origin)
c) Observe that the words match the reconstruction
d) conclude that the words under comparison are of IE origin because they match the reconstruction

This is a circular argument because our conclusions depend on the original assumption. It would be different if, for example, we had assumed that the words under examination were of different origin, belonging to different language families and then, during the reconstruction process, using the stated methods and the relevant data, we might have come to a conclusion different from the initial assumption. This type of circularity, in turn, stems (also) from the fact that the original choice of the languages to compare is, by its very nature, based on intuitive and often subjective judgments (these in turn being often influenced, or even dictated, by historical and socio-cultural factors). That the risk in question is a real, concrete one, and not just a theoretical possibility, can be shown with an equally concrete example: the issue of the well known correlations existing between Hungarian (traditionally classified as Uralic) and Turkic (traditionally classified as Altaic). These correlations are explained as contact-induced, the result of

long lasting, intense contacts between the early 'magyar' tribes and Turkic tribes (these contacts would have taken place after the magyar-s split off from the rest of the Uralic proto-community). However, there is no independent evidence to support the thesis of this centuries-long 'symbiosis' (as is usually referred to in the literature) and, therefore, of the ensuing borrowing, either from history, or archaeology, or anthropology (let alone the thesis of the splitting off of the magyar-s from the Uralic proto-community). On the contrary, the symbiosis in question is assumed *only* on the basis of the 'intense borrowing'. Thus, the whole argument represents a classical example of vicious circle. And, in fact, several scholars nowadays call into question the validity of both the Uralic (see for example Marcantonio [2002]) and the Altaic theory (see for example Unger [1990]). A similar situation is encountered within IE too, with regard to the well known lexical, phonological and structural similarities identified as existing between Sanskrit and the other (non-IE) languages that constitute the complex 'South Asian linguistic area' (see Masica (1976) as well as the 'Aryan Debate', as discussed by Bryant and Annamalai & Steever in this volume). These identified similarities have been traditionally interpreted by most scholars as 'clear' instances of borrowing, despite the fact that it is often difficult in this area (too) to tell apart similarities attributable to genetic inheritance, or to borrowing, or even to chance resemblances, as pointed out by some scholars; see for example Hock (1996: 36ff.).

Even if this 'intrinsic circularity' (as I would like to call it) were not considered to represent a methodological difficulty, there is still embedded into the comparative method what could be called the 'every-day practice circularity', which has been effectively described by Morpurgo Davies (1998: 254)[3] as follows:

> We can say that the sound laws have no exceptions only because when we find an exception we eliminate it saying that there has been analogical interference. On the other hand, we also say that

> the only way in which we can prove that a form [...] is analogical is by pointing out that otherwise it would be an exception to the sound laws

Other typical means of 'explaining away' exceptions and counter evidence are: assuming borrowing (even from unknown, extinct languages /dialects), re-arranging the stated sequence of rules in a different order, postulating a (/another) laryngeal segment, etc. It is true to say that each of these 'adjustable parameters' does reflect a plausible, genuine linguistic process. However, the problem is that the overall cumulative effect of many adjustable parameters added to the definition of a given law may endanger the 'cumulative effect', i.e., the 'statistical significance' any established law or tendency should possess. Thus, there is the danger that the explanatory system ceases to be 'explanatory' and becomes an 'ad-hoc' system, difficult to verify or, conversely, falsify. In addition to this, one may observe that the task of the comparative method is that to reject words that do not meet the stated sound-rules, therefore it can be no surprise that (at least in principle) one is left with only those words that (supposedly) do meet them. This way of proceeding amounts to a licence to reject counter evidence, or, to put it in other words, it introduces a systematic bias in the data which needs to be filtered out (as we shall see, this bias can be filtered out using quantitative methods of analysis).

1. 2. The impact of circularity on reconstruction

The circularity issue (to my knowledge) has hardly ever been the object of a targeted and comprehensive investigation, whose purpose would be that of assessing the statistical significance of the traditionally established IE comparative corpus (but see Brady & Marcantonio (2003), Marcantonio (2002) and [2003/5]). There may be several reasons for this. It has been often claimed that the comparative method is not a 'heuristic method', and therefore it can only be used 'after' a language family has been established

(through other means), to trace back sound changes and related correspondences (see for example Nichols [1996]). This being the case, it does not matter how many exceptional sound changes, how many 'false matches' there may be within a given comparative corpus. Besides, a certain amount of counter-evidence would not, in any case, be sufficient to call into question a theory as well established as the IE theory. For example, Ringe (1995: 60) argues that: "irregularity (in suitably small number) can be tolerated in working with languages whose relationship has already been established beyond doubt, but they are potentially fatal weaknesses in an attempt to prove a doubtful relationship". Similarly, It has been claimed that there are linguistic areas for which we 'know' that the languages are related, but whose relatedness cannot be demonstrate by using the logic of the comparative method, mainly because the 'correspondences' are, simply, not regular. This is for example the case of the languages of South-eastern New Caledonia (Grace 1990 & 1996), whose investigation appears to show that the languages are 'obviously' genetically related, although the method of comparison is hardly applicable. Following the same line of reasoning, other scholars have stated that the establishment of just a few cognates, or even of one single cognate, would be enough to demonstrate genetic relatedness. This is, for example, the position held by Harrison (2003: 217), who also observes that when the number of putative correspondence sets approaches a level that is not statistically significant (i.e., that might be attributable to chance) "the C[comparative] M[ethod] has ceased to work". Another reason for the lack of interest in assessing the impact of circularity on IE may lie in the fact that, in order to carry out this type of investigation, one needs to apply the appropriate methodology, i.e., quantitative, statistical methods of analysis. However, most linguists appear to believe that statistics is not a meaningful tool within historical linguistics. In fact, the task of the historical linguist would not be that of quantifying the evidence in favour or against a given model, a well established language

family, but that of finding those very factors that have triggered the encountered irregularities in order to justify them. In other words, counter-evidence is nothing but irregularities that we have not been able to justify (yet). This way of reasoning ties in well with the thesis that the comparative method is not heuristic, or that genetic relatedness cannot always be demonstrated, being at times simply 'compelling'. Thus, assessing the cumulative effect of sound laws does not help us in our understanding of language development and sound changes, and does not lead anywhere.

1. 3. The advantages of quantitative methods of analysis

In contrast to the points of view summarised above, there have been several linguists who have pointed out the advantages for historical linguistics to adopt (also) quantitative methods of analysis, and have themselves carried out several (types of) quantitative investigations. The major advantage would be that of relying on methods that are 'objective', being by their very nature 'replicable' and, therefore, truly verifiable or falsifiable; see for example Chrétien (1937); Ringe (1992, 1993, 1995, 1998 & 1999); Ringe et al. (2002); Mc Mahon & Mc Mahon (2003) and McMahon & McMahon[4] (2005). In particular, according to Ringe (1999: 213), it is fundamental that any observed similarity between two languages must be supported by a demonstration that the similarity in question could not be the result of sheer chance. Failing to provide a demonstration of this sort is "non-negotiable" for a linguistic discipline which purports itself to be scientific. On several occasions Ringe (quoted) also refers to what he defines: "the 'brute force' probabilistic analysis" that would correlate a very large number of pairs from IE languages, as against the poor statistical value correspondences which would correlate the various languages from other language families. Ringe (1995: 68-71) further observes that, since one cannot prove that languages are *not* related, those who propose specific correlations among languages have to provide "objective proof" of them, otherwise

"scientific historical linguistics is simply impossible". Similarly, McMahon & McMahon (2005:26-7) state that:

> The current difficulty comparativists face is our inability to test and demonstrate family relationships, so that these can either be proved beyond reasonable doubts, or refuted. If we cannot tell good results from bad ones in a formal and repeatable way, we cannot hope to distinguish good methods from bad ones either. [.....] The relative informality of the [comparative] method, and the lack of testability beyond the checks built into the method itself, mean we have to rely on the experience and integrity of individual practitioners to do so

Despite this kind of programmatic (and, in my opinion, correct) statements, hardly any quantitative investigation carried out thus far has actually measured the degree of statistical significance that (supposedly) correlates a very large number of correspondences from the IE languages. On the contrary, the statistical significance of the IE comparative corpus is taken for granted (but see Bird's (1982) and Ringe's (1995) analysis of Pokorny's dictionary[5]). The quantitative investigations in question have in fact centred on the following two tasks: a) to assess the soundness of the traditional, internal sub-branching of IE. For example, Ringe et al. (2002:81) attempt to recover the first order sub-grouping of the IE family using a new computational method, because through the traditional methods of historical linguistics no consensus has ever been reached "on how those ten robust subfamilies [of IE] are related to one another cladistically" (see also McMahon & McMahon [2005]). b) to compare the IE comparative corpus with that of the Nostratic /Eurasiatic macro-family (Ringe 1995 & 1998), in order to show how the latter lacks statistical significance and is not, therefore, a valid family. In this case the IE comparative corpus is simply used as a 'control case'.

Thus, whether implicitly assumed or explicitly stated, whether considered to be a relevant or an irrelevant factor, the positive statistical significance of the IE corpus appears to be taken for granted, and this is the point where my personal opinion differs from that of most linguists'. In other words, I believe that:

a) Assessing the statistical significance of the comparative corpus of any language family (including IE), does actually matter. In my opinion, when the number of putative cognate sets approaches a level that is not statistically significant (to use Harrisons' words) this does not mean that "the C[omparative] M[ethod] has ceased to work"; on the contrary, it means that we are facing clear instances of evidence counter to the predictions of the model;
b) Even if one did not agree with this stand of mine (and, as we have seen, many historical linguists in fact do not), I believe it is nevertheless a useful exercise to investigate the statistical significance of the IE comparative corpus. It would in fact be desirable to obtain an empirically assessed, objective picture of the status of IE with regard to, for example, those language families that hardly contain any proper correspondences (see discussion above; par. **1.2.**);
c) If one accepts what is stated in points (a) and (b) above, then the statistical significance of IE has to be properly and empirically verified, instead of being simply assumed, and this verification can only be achieved through the help of quantitative, replicable types of analyses.

Bearing these general issues in mind, I shall examine a significant sample of the conventionally established IE phonological / lexical and morpho-phonological comparative corpora. In practice, I shall attempt to asses the ratio between the number of rules (and related 'adjustable parameters') needed to establish a given correspondence and the overall number of correspondences those

rules are capable of accounting for. Once the picture on this aspect of IE studies is clear one can draw his/her own conclusions on whether the attained results have any bearing on the validity, or otherwise, of the theory as a whole, according to his/her own stand on the general issues discussed above and in the *Introduction*.

1. 4. The quantitative analysis: Summary of the methods and results

The data examined in this investigation are taken from a well defined source, the **LIV** dictionary of IE verbal roots (Rix 1998). This is a modern, (supposedly) much more rigorous dictionary than the previous existing ones, a dictionary that represents the current, widely accepted way on the procedures of reconstruction of IE roots and stems. My investigation involves simply counting and categorising all the reconstructed verbal roots present in **LIV**[6], as well as any other relevant element associated with the roots in question, such as meaning(s), morpho-phonological alternations, presence *vs* absence of laryngeal segments, etc. I shall then simply map the results of the counting into standard types of graphs. No sophisticated statistical analysis and /or computer modelling of the type used by Ringe (quoted) or McMahon & McMahon (quoted) are in fact used here. The quantitative investigation of the verbal roots is designed to quantify the following main instances of circularity, as indicated by the reported results:
Result 1: 32% of the roots are reconstructed on the basis of data drawn from one single branch of the IE family. Thus, the evidence suggests that these roots may be of local or other origin. However, the assumption is made that they are of IE origin and reconstructions are proposed; it is then self-evidently circular to conclude that these words are of IE origin just because of the establishment of a reconstruction.
Result 2: 34% of the roots are reconstructed on the basis of data drawn from two language branches only. As is known, reconstructions of this sort are not safe, because the possibility of

chance resemblance or borrowing rates quite high in these circumstances. In spite of this, the assumption is made that these roots are of IE origin, and reconstructions are proposed. However, once again, it would be circular to conclude that these words are of IE origin because of the existence of reconstructions.

Result 3: Most reconstructions of the verbal roots consist not just of a single reconstruction, but of a set of alternative 'sub-reconstructions' – a set of alternative starting-points which are called 'alternations'. This is justified by the observation that morpho-phonological alternations are indeed one of the features characteristic of (several) IE languages. Although this is undoubtedly the case, if not used carefully, this procedure can make the method circular. To take an extreme example, if an attested language **A** uses the word *dog* whilst another language **B** uses the word *cat* to refer to what is supposed to be the same or similar reference, then if one assumes (a priori) that the words belong to genetically related languages, one might reconstruct the proto-word **cat* ~ **dog* in order to justify the actual difference in sound (and often also in meaning). In the analysis below I'll show that the number of alternative 'sub-reconstructions' reconstructed for each verbal root typically increases -- and increases linearly -- with the increase of the number of the language branches that support the root. This is exactly what one would expect if the **cat* ~ **dog* fallacy were being used. No alternative explanation for this clear pattern in the data can really be found. Thus, reconstructions of this sort are based on the assumption that the words are related, and so it would be circular to conclude that in fact they are so related, just because of the establishment of 'reconstructed' – as against to actually 'attested' – morpho-phonological alternations.

As mentioned, the object of my investigation is the phonological (/ lexical) correspondences. This is not just because **LIV** deals with lexical correspondences, but because at the phonological level there is at least one clear (although questionable) criterion against which scholars can verify the

attained results: the regularity principle. I shall not investigate the conventionally established morphological correspondences because, at this level of language, it is not clear which guiding principles and criteria can be adopted in order to establish a match, if not an intuitive, 'naked-eye' observation of (a certain degree of) similarity among morphological elements. In other words, it is not clear how to validate the morphological model against the evidence, since there are not clearly specified laws and there is much greater reliance on processes such as analogical leveling, whose conditions of operation are not defined using predictive and verifiable criteria. However, I have examined a significant sample of morpho-phonological correlations, which can be analyzed through quantitative methods because of the (supposedly) rigorous correlation existing between sound alternation and morphological information. As we shall see, the results of this investigation show that the phonological/lexical corpus and at least one specific morpho-phonological corpus assembled in **LIV** definitively lack statistical significance, as a consequence of which it is reasonable to assume that most of the correspondences there assembled are 'chance resemblances'.

The possibility that the traditionally established comparative corpus of IE (like that of any other language family) might contain a (more or less) high number of chance resemblances should not come as a surprise. For example, Hock (1993 & 1994) considers the issue of chance resemblances to be an "especially troublesome" issue. He has assembled long lists of chance resemblances among several IE and Dravidian languages[7], just to observe that (1996: 37): "As the continuous controversy over remote linguistic relationships shows, there is no generally accepted answer to the question: What is the chance of similarities being accidental?". Well, I hope the analysis carried out below will enable us to get a fairly faithful picture of the level of chance similarities embedded within IE.

2. The statistical significance of the Indo-European comparative corpus assembled in LIV

2. 1. Counting the verbal roots

Counting all the verbal roots listed in **LIV** one observes that 200 of the 683 'safely reconstructed' roots ('safely' according to the dictionary itself), are reconstructed on the basis of only one daughter language or language branch ('witness' to the reconstruction). In the absence of any comparison with another language branch, it is hard to discern the basis of the claim that the roots descended from ancient IE, rather than from some other source, much less that they have been 'safely' reconstructed as Indo-European. In fact, if a word is present in only one language group then this strongly suggests quite the opposite: the word is lf local origin. Likewise, an equally high number of roots (34%, as mentioned) have been reconstructed on the basis of only two daughter language branches as 'witnesses'. There is an unacceptably high likelihood of false matches if one operates with only two 'witnesses' to support the reconstruction, because it is too easy or tempting to pick out similar forms that may or may not be truly related. This observation is generally accepted among linguists; for example, Meillet (1934: 41) proposes the 'three-witnesses criterion', whilst Greenberg (2005) suggests four languages as a sufficient number to establish a historical relationship.

This simple counting shows that, of the **LIV** corpus, 66% of the roots cannot be relied upon because they are reconstructed on the basis of just one or two languages / branches witnesses only. They lack therefore statistical significance, and it is hard to imagine on which quantitative basis they can be considered as reflections of genuine IE correlations. As a matter of fact, Campbell (1998:115), obviously accepting the three-witnesses criterion as relevant, criticizes the 'Nostratic' comparative corpus as assembled by Illič-Svityč (1971-84) for the fact that 35 % of the

cognate sets have been established on the basis of only two member families. On the basis of this and other methodological flaws Campbell rejects the Nostratic hypothesis as not scientifically proven (correctly, in my opinion). Thus, on the same basis, for the sake of consistency, one has to reject 66% of the verbal roots as reconstructed in **LIV**. Discounting those unreliable roots whose reconstruction does not meet the three-witnesses criterion, this leaves us with 34% of the roots whose reconstruction is supported by three or more witnesses and are therefore worth examining as the basis of a genuine IE comparative corpus. Of this 34%, nearly half have been reconstructed using laryngeal segments, whose status as phonetic/phonological IE segments is still disputed, despite the fact that the laryngeal theory is now widely accepted. This leaves 126 roots (i.e. 18% of the total), that could form evidence, in principle, for genuine linguistic correlations, because their reconstruction meets the three-witnesses criterion and do not make recourse to laryngeal segments.

The next step will be to try and count the number of linguistic rules that have been used to reconstruct the **LIV** corpus. Unfortunately **LIV** does not state which rules are used in its reconstruction, and therefore one must look at the general literature and assume that these rules have been adopted in the dictionary. According to Collinge (1985), there are 54 'major laws' and 11 'major tendencies' within IE, giving a total of 65 general rules. In addition, for each rule there are a variable number of 'contextual specifications' or 'adjustable parameters'. The precise number is very difficult to ascertain, because, of course, this number depends on single interpretations of a given Law (of which often there are several): those interpretations that admit fewer exceptions tend to contain more adjustable parameters, since these are needed to justify the exceptions in question and transform them into 'apparent exceptions'. Brady & Marcantonio (2003) have counted 165 to 202, general (as opposed to language-

specific) such adjustable parameters for all the Laws (not major tendencies) listed in Collinge (1985). This estimated number of contextual specifications/adjustable parameters represents an under-estimate, since the contextual, 'language-specific' specifications that are evidently used in **LIV** (in the form of 'explanatory notes') to get the expected reflex in all the relevant languages, but have not been mentioned by Collinge while merely stating the general Law, have not been counted.

Whatever the overall number of Laws plus their adjustable parameters might be, their number is at least comparable with, and almost certainly higher than, 126, the number of the roots that meet the three-witnesses criterion and do not contain laryngeals. These roots have been called in Brady & Marcantonio (2003) as 'lawful' – just to pick up a name – that is, they have been established according to the conventional Laws, without making recourse to laryngeal segments. The diagram below (from Brady & Marcantonio 2003) illustrates this situation: there is only a very small number of roots that are widely represented, being attested in 6 or more languages / language branches; this number becomes even smaller if one does not count the roots containing laryngeal segments (or if one does not accept the particular version of the laryngeal theory adopted by **LIV**; see below):

Figure I: the Indo-European verbal roots and the 'three-witnesses' criterion

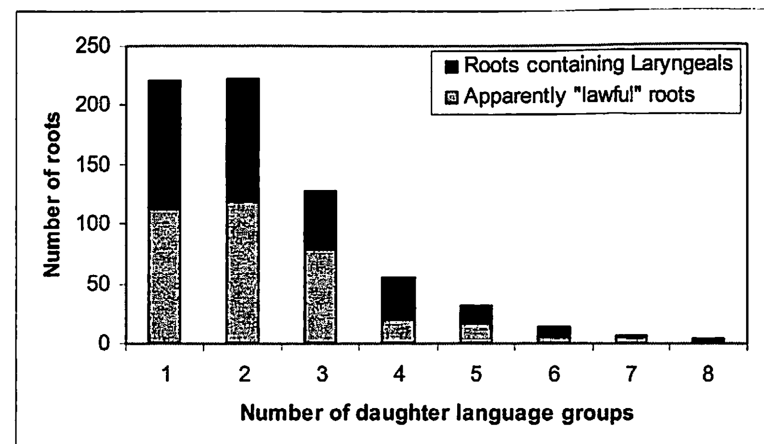


Figure I: The number of roots that support the Laws of Indo-European. Only 126 of the 683 reported roots (18%) meet the 'three witnesses' significance criterion and do not contain laryngeal segments.

2. 2. Are the Indo-European morpho-phonemic alternations genuine or artefacts?

2. 2. 1. We shall see in this section that there is evidence of a strong bias in the reconstruction of the roots themselves in the area of vowel alternation (*Ablaut*), and associated morpho-phonology. This bias has helped researchers to find correlations where, according to my analysis, they are not justified and may be an artefact of the method of analysis. As we shall see below, the bias is betrayed by the fact that the roots whose reconstruction has been established on the basis of a large number of IE languages / language groups tend also to display a large number of morpho-phonemic alternations.

Vowel alternations, both in nouns and verbs, are widely claimed to be a distinctive feature of IE, and many of the vowel alternations present in the modern languages have, supposedly, been inherited from alternations already existing in the proto-

language (see the chapter by Carruba in this volume for a detailed account of this phenomenon). In particular, among the inherited alternations in the verbs there are claimed to be morpho-phonemic alternations of the radical vowels ('radical *Ablaut*'), where different grades are associated with different tenses (at the singular active within a given declension). Thus, most root reconstructions typically consist not just of a single reconstructed root, but of a set of alternative reconstructions, postulated in order to represent (also) the various morpho-phonemic alternations supposedly already existing in the proto-language. For example, according to Szemerényi (1996: 91), the set of the basic *Ablaut* variations of the disyllabic root **genē-* 'to bear, be born, beget' is the following:

**genə-*, **gnē-*, **gn̥-* (for the *e*-grade series); **gonə-*, **gnō-* (for the *o*-grade series) and **gēnə-*, **gōnə-* (for the long grade series, in first syllable only). Since *schwa* can be lost before vowels, also forms of the type **gen-*, **gon-* and **gēn-* can occur

Although gradation is a common phenomenon in languages, it is intuitively clear that, with so many starting-point reconstructions, it can be a priori quite easy to match almost any form found across a given linguistic area. Indeed, for example, it has been necessary to postulate up to ten alternative reconstructions for **genē-* in order to account for the few attested and spottily distributed forms. Actually, in this case the process has almost been taken to the extreme, since the number of the reconstructed alternations (10) is close to the number of the attested ones (there are 16 different forms from only 4 languages: Sanskrit, Latin, Greek and Germanic). In other words, postulating so many gradations at the level of the proto-language makes it an easy job to 'prove' that the specific, attested patterns of alternation (rather than, say, the general phenomenon of gradation) have been inherited from the supposed proto-language[8]. In layman's terms, we are close to the hypothetical extreme discussed above, where we have two attested

words, *dog* and *cat*, and reconstruct that they are connected via the IE form **cat* ~ **dog*. We shall examine this in more detail below.

Let us now discuss in detail the verbal roots as reconstructed in **LIV**. Each reconstruction consists of two or more alternative reconstructions (alternative forms), which I shall call 'variants'. These variants consist, mainly, of vowel alternations / morpho-phonemic variations, although often they also contain affixes (infixes/suffixes). These affixes – called by **LIV** 'primary affixes' – enlarge the original root and are claimed to convey aspectual information and /or to modify the original meaning of the bare root. This is certainly a common, genuine linguistic phenomenon, although it is self-evident that many roots do display a bewildering array of different affixes whose origin and function cannot be traced back and whose reconstruction is therefore hard to justify (for a discussion of this issue see Clackson [2007:151 ff.]). Like the root, affixes too may alternate. In a word, by 'variant' I mean each different alternating form – consisting of *Ablaut* only or *Ablaut* + primary affix – that, together with the other, similar forms in the set, contributes to make a reconstructed root (and, of course, the various, reconstructed stems derived from it). These variants are regarded as different aspects of a single reconstruction, and many of them contain laryngeal segments, either in the root or in the affixes, or both (according to the version of laryngeal theory adopted by **LIV**). Let us see a couple of examples: the root ***deh$_3$** 'to give' (***dō-**, according to pre-laryngeal notation; **LIV** 1998:89; 2001:105-6) and the root ***dek'** 'to take, perceive/notice' (**LIV** 1998:93-5; 2001: 109-11):

***deh$_3$** 'to give'

1) Aorist **déh$_3$- / dh$_3$-* 2) Present **dé-doh$_3$ /dh$_3$-*
3) Desiderative[9] ?**di-dh$_3$-sé-*

***dek'** 'to take, notice'

1) Aorist _*dék'-/d_ek'-_	2a) Present _*dḗk'-/dék'-_
2b) Present _*d_ek'-néu̯/nu-_	3) Perfect _*de-dók'/d_ek'-_
4) Causative _*dok'-éi̯e-_	5) Desiderative: _*di-dk'-sé-_
6) Intensive _*dék'-dok'/d_ek'-_	7) Essive: _*d_ek'-h_1u̯é-_

As one can see, the root **_*deh₃_** consists of a set of five alternative forms, five variants in total, across 3 tenses / moods: two variants associated with aorist, two with present and only one with desiderative, the latter being enlarged with the suffix _-sé-_. Similarly, the root **_*dek'_** consists of a set of 13 variants across seven tenses/ moods: two variants associated with aorist, two with type (a) and two with type (b) present[10]; two variants associated with perfect; one with causative and one with desiderative; two again associated with intensive, and one with essive. Type (2b) present, causative and desiderative are also enlarged with affixes. The suffix _-néu̯/nu-_, which enlarges the present tense of type (b) in the root **_*dek'_**, also, in turn, displays alternation. The fact that there is not just one single form / variant, but a sub-set of two alternating variants associated with some tenses is justified by the observation that these tenses tend to display alternations also 'internally', within their declension (according to conventional reconstructions). Typically, an alternation between full grade in the singular active of the indicative/injunctive _vs_ reduced or Ø-grade elsewhere. In other words, the overall set of variants of each root typically consists of what one could call 'horizontal' and 'internal' variants, that is, respectively: alternating forms associated with different tenses (such as present _vs_ perfect), and alternating forms associated with the internal paradigm (such as singular _vs_ plural / dual within a given declension).

As an example of a specific (horizontal) morpho-phonemic alternation let us consider the present _vs_ perfect _vs_ aorist alternation in Greek, widely quoted as illustration of the

phenomenon. As conventionally stated, a full, _e_-grade is associated with the present tense, an _o_-grade is associated with the perfect tense, whilst the aorist is associated with the Ø-grade (at the singular indicative/ injunctive active), as illustrated by the Greek[11] verbs 'to leave' and 'to see' below:

Present	Perfect	Aorist
λείπω 'leave'	lέ-λοιπ-α	ἔ -λιπ-ον
δέρκο-μαι 'see'	δέ-δορ-κα	ἔ-δραk-ον (* < dr̥k-)

Morpho-phonemic alternations of this sort, if substantiated by a good amount of items of evidence (whatever 'good' might mean) across the IE languages, could be indeed a strong clue of genetic inheritance, because a widely attested, constant association between a given radical vowel grade and a given morphological information can hardly be the result of chance resemblance, or borrowing. However, the reality is that the evidence for postulating so many variants in the reconstruction of a single root (including many types of present and several enlarging, primary affixes) appears to be a fabrication of the method because of the following facts: A) as anticipated, roots established on the basis of a high(er) number of IE daughter languages / branches have been reconstructed with a high(er) number of alternative variants, this being an indication of a strong bias (see par. **2. 2. 2.** below); B) a detailed investigation of one specific morpho-phonemic alternation, the present _vs_ perfect, has revealed that this alternation is actually poorly attested, thus confirming the suspicion that most of the postulated, reconstructed alternations may well be artefacts of the method of analysis (see par. **2. 2. 3.** below).

2. 2. 2. Let us then start with the issue raised in point (A) above, that is: the correlation between the number of languages / language groups that support a given reconstructed root and the number of variants reconstructed for that very root. Here the crucial factor is

that this correlation is quite strong, and the constant of proportionality is very close to 'one'. This suggests that, on average, one variant (/alternative form), has been added to the root each time a new daughter language (branch) is claimed as being supporting evidence for establishing that root. There are several hundred such roots in **LIV**. For each verbal root, I have counted the number of reconstructed variants (**V**) associated with the root, and the number of daughter languages (language groups/branches) (**G**) in which at least one (supposedly) derived verbal item is attested. These two values (**V** and **G**) should be independent of one another if the correlations among the assumed cognates (and related distribution of radical *Ablaut*) are to be considered genuine. On the other hand, if the reconstructions were due to chance resemblances, one would expect that the two values **V** and **G** would be positively correlated, because the higher is the number of variants postulated, the higher is the likelihood of establishing a match, or better, a false match among the language groups. This is indeed what my counting has revealed. As shown in the graph in figure **II**, the average number of variants is strongly correlated with the number of daughter language groups accounted for by the root. In fact, the data fits extremely well to the formula **G** = 1.6 + **L** for the range 1 to 5 groups:

Figure II: the alternative reconstructed forms and meanings ('variants') in each root:

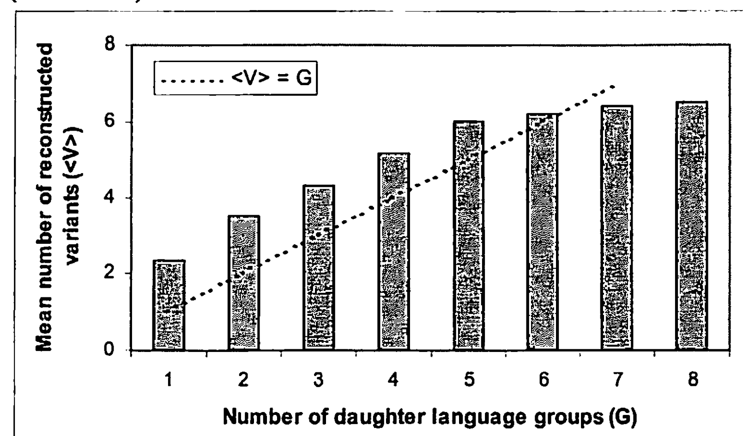


Figure **II**: The number of alternative reconstructed forms and meanings ('variants') in each root. For the 96% of the roots with up to five claimed daughter languages /language groups, the graph has almost the same slope as the dotted line. This indicates, on average, that one new variant has been 'appended' to the reconstructed root with each newly claimed daughter language group. There is, however, a departure from linearity where **G** > **5** (see figure **V** below).

Thus, this clear correlation, with a proportionality factor of very close to '1', strongly suggests that, on average, one new variant has been 'appended' to the reconstruction with each newly claimed daughter language group. In turn this suggests that the majority of the 'cognates' reported in this corpus are artefacts of the method of reconstruction. There is, nevertheless, a departure from this linearity, whereby **G** > **5** (see figure **V** below). In other words, there is a small number of roots – 4% – whose reconstruction is supported by 6 or more daughter language groups and that are, therefore, potentially (statistically) significant. I shall discuss this aspect below, attempting to find a possible explanation for this departure from the observed linearity.

Coming back to the issue of the linear relationship between the two values (**G** and **V**), linguists have already asked themselves why there are many fewer alternations in the attested languages than in the reconstructions. Some scholars (such as Szemerényi (1996:83-93); Sihler (1995: 109-111) and others) assume that different languages may inherit different variants from the multiplicity of true, original IE variants of a root. Thus, it should not be surprising that only a few and /or different grades (out of a complete series) are found scattered across different languages. In other words, each new language group claimed to support the root under investigation would uncover more variants from the original, complete set. Other scholars, as well as the **LIV** dictionary, assume (also) a wide-spread process of analogical leveling in the daughter languages (a similar interpretation is offered by Carruba in this volume, although Carruba admits that the IE *Ablaut* has not yet been fully understood and explained). As an example of how one particular variant out of a given set of reconstructed variants may be attested by one language or two, another variant by a different language, a third alternation by yet another, different language, and so on, let us consider again the root ***dek'*** (1998:93-4; 2001:109-11). The 7 reconstructed (horizontal) variants, according to **LIV**, are distributed as follows:

1) the two forms of present (a) **dḗk'-/dék'-* and (b)**d_ek'-néu̯ /nu-* are attested in Indo-Iranian and Greek;
2) the perfect **de-dók'/d_ek'-* is attested in Vedic, Greek and Latin
3) the aorist **dék'-/d_ek'-* is attested in Greek and Armenian
4) the causative: **dok'-éi̯e-* is attested in Hittite, Greek and Latin
5) the desiderative: **di-dk'-sé-* is attested in Vedic and Latin
6) the intensive: **dék'-dok'/d_ek'-* is attested in Greek
7) the essive: **d_ek'-h₁i̯é-* is attested in Latin

These conventional explanations are certainly plausible, in principle. However, the thesis of the analogical levelling turns out to be an ad-hoc explanation, since there is no independent evidence for this. As to the thesis that more languages uncover more variants, combining this explanation with the fact that the proportionality is almost exactly '1' leads to an important conclusion: almost every reconstruction is based on a different starting point reconstruction (the extreme '**dog* ~ **cat*' situation). Furthermore, this explanation is inconsistent with the fact that the 'first' language group is special in that it uncovers more variants than subsequent groups, as is evident from the graph (Fig. **II** above). In practice, the evidence suggests that the following process has taken place. A reconstruction is initially introduced with about 2.3 variants on average, based on some of the variants actually attested in some daughter languages (mainly Sanskrit and, to a lesser extent, Greek), from which comparisons typically start. If, while searching for other language groups in support of the reconstructed root, the evidence does not quite match the initially proposed reconstruction, then one new variant can be appended to accommodate it. Thus the linear relationship between **V** (variants) and **G** (language group), and its slope of (almost) 'unity', is the statistical result of this process of introducing a large number of alternative reconstructions, which are artefacts of the method of analysis. In other words, here there is clear evidence of a bias, which appears to have arisen because linguists, analysing the data, knew a priori which words are 'supposed' to share a common ancestor.

2. 2. 3. Let us now discuss the issue raised in point (B), i.e.: the distribution of the *e vs o* (present *vs* perfect) alternation, one of the major sources of variants of a root within IE. Accordingly, in **LIV** 259 roots display this alternation and 114 of these are considered to be safe cases. However, a close scrutiny of the actual evidence across the IE area reveals that this alternation is rather poorly

attested. In fact, only the Indo-Iranian languages and Greek document a clear category of perfect with *o*-grade (strong form) in the singular indicative active, alternating with *Ø*-grade (weak form) elsewhere, with related position of the accent on the radical syllable in the strong forms and on the desinential syllable in the weak forms. The situation is much more complex and unclear in the other languages, including the Germanic languages and, even more, Latin (see below). Furthermore, even in Indo-Iranian and Greek the evidence is not that clear-cut after all. In fact, in Indo-Iranian the -*o*-grade is attested indirectly, due to the merger[12] of short /e/ and /o/. In Greek, the forms of perfect presenting *o*-grade *vs* *Ø*-grade are only a small number (about a dozen verbs) that are, however, considered to be very old; see the type: Sing. οἶδα *vs* Plu. ἴδμεν [13], corresponding to Skt. *vḗda* vs *vidmá* (< **woid*; 'I /we have seen', therefore 'I /we know'). Basically, in Greek the internal alternation (in the appropriate declensions) is found only in the so-called a-thematic formations[14], a small minority of cases, whilst the absence of internal alternation is the norm. This contradiction between the predictions of the model and the actual data is 'explained away' through typical 'rescuing' procedures, one of which is analogy (as mentioned). As Sihler (1995:109) puts it: "In Greek the inherited pattern [of *Ablaut*] have been analogically extended, leveled and otherwise confused". As to Germanic and Latin, the situation is as follows. In Germanic, the present *vs* perfect, horizontal alternation, as well as the internal alternation between full and *Ø*-grade is attested for sure[15] in one type of perfect only, the so-called '*präteritopräsentia*' as shown in Goth. *wait* ~ *witum* (*wissa, wiss*), corresponding to the above mentioned οἶδα ~ ἴδμεν and *vḗda* ~ *vidmá* (there are about ten examples of this type; for more details see Prokosch (1938: 188 ff.); Di Giovine (1996 II:139 ff.) and Clackson [2007:120 ff.]). In Latin, the overall *Ablaut* system, including that of the perfect, admittedly, is quite different from the Indo-Iranian and Greek one, although opinions may vary on how to interpret such a wide

difference, such a "ruin", as Sihler (1995:109) defines it. For example, Di Giovine (1996 II: 140-1) stresses the basic independence of the Latin perfect system from the IE one, because there are no safe instances either of *o*-grade in the radical syllable or of alternation between 'strong' ~ 'weak' grades within the declension:

> Nella flessione del *perfectum* latino [...] non si evidenziano *sicuri* esempi di continuazione di un grado vocalico *-*ŏ*- nella sillaba radicale [...] si deve anche osservare che non esistono esempi di alternanza tra forme forti e forme deboli all'interno della flessione del *perfectum*[16]

On the contrary, Sihler (*ibidem*) argues in favour of the basic IE character of the Latin perfect, although a series of innovations would have almost completely "obscured" its original shape (see the chapter by Di Giovine in this volume on the perfect and the IE verbal system in general)

In order to quantify this situation, I counted those roots whose reconstructions include both the present and perfect tense, *e* ~ *o* alternation. In seventeen of these reconstructed roots, none of the daughter language groups actually attest the alternation in question. In other words, here the alternation has been reconstructed in the absence of any actual evidence, using the assumptions of the model alone. Further, in the vast majority of the roots (73), only 1 daughter language group shows the alternation. This evidence, if interpreted without the bias of the model, clearly suggests that the alternation is simply of local origin. Figure **III** below shows the full results of the counting. The present ~ perfect alternation is in fact supported by only 8 roots that meet the 'three witnesses' criterion (See Table **II** below for the list of these roots). Examining these roots in more detail, one finds that even these do not provide straightforward evidence for

the alternation because: a) 6 of them contain laryngeal segments; b) 7 of them include Sanskrit, where *e* and *o* merge into *a*:

Figure III: The present ~ perfect alternation

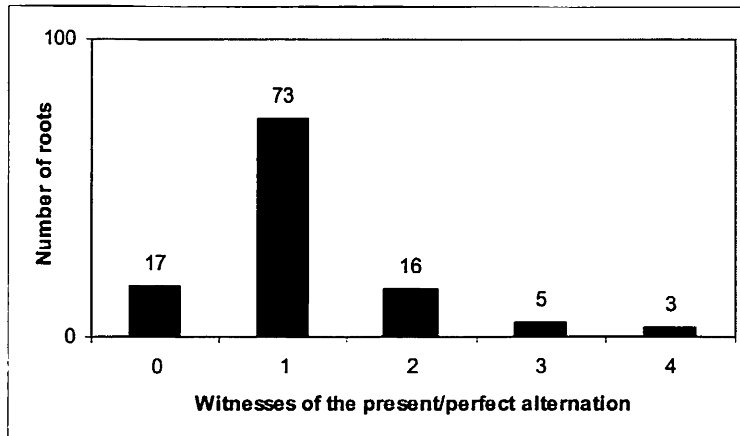


Figure **III**: Number of roots supporting the present ~ perfect alternation. This graph shows those roots whose reconstructions include both present and perfect tense, distributed by the number of supporting language groups, whereby the cognates are reported with both these tenses ("witnesses of the present ~ perfect alternation").

To take a concrete example of how an attested verbal form (including perfects or presents) in a given language can be associated to just one out of the set of variants of its reconstructed root, without necessarily displaying any horizontal and / or internal gradation, let us consider the root *$*b^h\underset{\circ}{u}eh_2$-* 'to grow, originate / be created, become, be' (**LIV**: 83-5), whose present and perfect are, respectively:

Pres. $*b^h\acute{e}\underset{\circ}{u}h_2$*-e-* ~ $*b^huh_2$*-ié-* *vs* Perf. $*b^he$-$b^h\underset{\circ}{u}\acute{o}h_2$ ~ b^huh_2-

This is one of the few potentially 'statistically good' roots, in the sense that the present, the perfect and also the aorist (whose

alternative forms are $*b^h\underset{\circ}{u}\acute{e}h_2$ - ~ b^huh_2-) are actually attested in at least three languages (see lexical item (6) in Table **II** below). Nevertheless, most attested verbal forms derive from just one of the reconstructed variants of the assumed set of alternations. For example, associated with the perfect there are the following, non-alternating forms: (**LIV** 1998: 83; 2002: 98):

- Sanskrit (Vedic): Perf. *ba-bhū́-va* 'he/she[17] is (has become)';
- Greek: Perf. *πεφύασι* 'they are / have grown';
- Oscan: Perf. *fufens* 'were';
- Old Norse: Pret. *bjó* 'lived'

The Oscan imperfect form *fufans* 'were' would also derive from just one variant of the perfect of the root, this being the full grade $*b^he$-$b^h\underset{\circ}{u}eh_2$ (according to footnote (24) of **LIV** 1998; but see also Sihler (1995: 554-5), according to whom there are uncertainties regarding the precise shape of the original stem). Similarly, associated with the aorist are the following, non-alternating forms, all derived from the single variant b^huh_2-:

- Vedic: Aor. *á-bhū-t* 'he / she has become';
- Greek: Aor. *ἔ-φῦ* 'he /she grew, became, originated';
- Old Latin: Conj. *fu-ās* 'you should be';
- Latin: Perf. *fūī* 'I was';
- Old Lithuanian: *bit*(*i*) 'he was';
- Old Church Slavic: *by* 'he was';
- Modern English *be* (a finite stem in Old English *béom*, *bist* etc.; see Sihler [1995:552]).

The only alternating, attested forms available, derived from both internal variants of the aorist, are the Old Irish ones: 3rd Sing. *boí* 'was' alternating with 3rd Plu. *-bátar* 'were'. Furthermore, as one can see, not necessarily an attested tense (/mood) derives from the

respective reconstructed one, as shown by the case of the Latin perfect *fūī*, and, ultimately, the present forms of Modern and Old English, all derived from one variant of the reconstructed aorist. This type of mismatch between tense and /or mood categories are encountered frequently in the process of reconstruction, and in **LIV**. Although we are dealing with a genuine and common phenomenon, the fact remains that these mismatches between the reconstructed and the attested verbal categories further increase the flexibility of the explanatory system, especially in connection with the variety of meanings admitted for most reconstructed roots. In the specific case of *$*b^h$u̯eh$_2$*, the root is associated with several basic meanings, six variants and one laryngeal segment (see Table **II** below). All these elements, coupled with the fact that this root alone has in the Rigveda several thousand forms (of which there are over twenty instances of the perfect stem) clearly grossly increase the possibility of 'picking up' the form suitable to obtain the desired match. To conclude this paragraph, two final observations: a) quite often, the single languages for which one more variant has been 'appended' to the reconstruction of a root are Tocharian and Hittite; b) the number of variants that constitute the set of alternative reconstructions for a given root in **LIV** can rise up to 13.

From the facts and analyses reported above one can draw the conclusion that the great majority of the morpho-phonemic alternations listed in **LIV** are not real, inherited alternations, but just artefacts, a fabrication of the method of analysis.

2. 3. Filtering out the artefacts

One can make a very rough estimate of the effect of filtering out the artefacts described above. If a root is reconstructed with more variants than language groups/branches to support them, then it should be discounted as having no significance. Putting this more mathematically, one can define the 'Intrinsic Significance' (**S**) value of a root as the amount by which the number of language

groups (**G**) exceeds the number of variants (**V**), i.e.: **S = G - V**. Roots with a 'negative' Intrinsic Significance value are those displaying more variants than language groups/branches – with the appropriate, attested morpho-phonemic alternation – to support them, and therefore are of doubtful significance.

Of the 357 roots listed in **LIV** whose reconstruction has been achieved without the aid of the laryngeal theory, 39 (11%) have a 'positive' Intrinsic Significance value, i.e., their reconstruction contains more languages / language branches as witnesses of the reconstruction than variants. Of these, 29 roots (7%) are reconstructed on the basis of three or more language branches witnesses; therefore these may be regarded as qualified candidates to be statistically significant correspondences and to properly support the stated Laws of Indo-European. Thus, according to this analysis, there is indeed a bunch of reconstructed roots that might reflect genuine IE correlations, but its number is only a very small portion of the total: just 7% of the roots listed in **LIV** that do not contain laryngeal segments – much more than this of course (as shown in the graph below) if one accepts without reservations the laryngeal theory. Notice however that this positive evaluation does not take into account the number of Laws and adjustable parameters needed to reconstruct the roots in question:

Figure IV: the 'Intrinsic Significance' value of the Indo-European roots

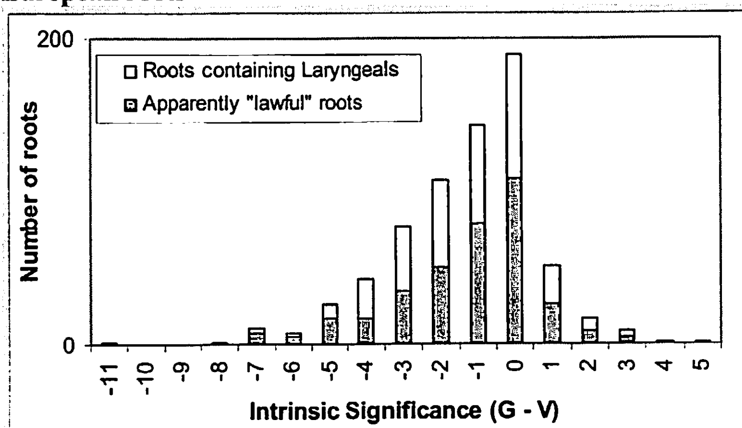


Figure IV: Number of roots distributed by their Intrinsic Significance value. Of the 357 'lawful' roots (not containing laryngeal segments), 39 (11%) have a positive Intrinsic Significance value (i.e. more language witnesses than variants). Of these, 29 (7%) have three or more language witnesses and may be regarded as qualified to represent real correspondences.

2. 4. The laryngeal theory: adding extra flexibility to the system?

Let us now come back to Figure **II** above (par. **2. 2. 2**), whose graph suggests that, on average, one new variant has been appended to the reconstruction with each newly claimed daughter language group, with the exception of a very small number of roots (4%). These roots have a higher average Intrinsic Significance than the others (**S** > **Ø** below the dotted line). Brady & Marcantonio (2003) have looked for a possible explanation for the departure from linearity of these 4% of roots whose reconstruction is supported by 6 or more daughter languages / language groups. A closer investigation has revealed that these significant roots are more likely to have been reconstructed using laryngeal segments. This in turn suggests that 'appending

laryngeal segments' increases the flexibility of the model, as an alternative to 'appending more variants'. The laryngeal theory is now accepted, and considered to be settled, by the great majority of scholars, and it certainly goes beyond the scope of this work to deal with the rights or wrongs of the theory. Here, as mentioned, I just limit myself to observe the effects the laryngeal theory (or at least the specific version of it adopted by **LIV**) may have on the flexibility and the statistical significance of the corpus under investigation.

Matching vowel alternations across the IE area (in both the mono-syllabic and by-syllabic roots) on the basis of actual evidence has always proven difficult, if not impossible, before the introduction of the laryngeal theory – as clearly illustrated in the chapter by Carruba in this volume. This is recognised also by **LIV** (p. 4) when it states that, without such a theory, "Morphostrukturen und Bildungsregeln des Urindogermanischen nicht verständlich sind". The dictionary adopts a system with four laryngeals: h_1, h_2, h_3, + a not well specified *H*. In particular, h_1, h_2, h_3 (according to **LIV**, p. 5), "Können fast jeden Platz einnehmen und an fast jeder Stelle zusätzlich zu den anderen Radikalen stehen". **LIV** does not divulge any information about the number of Laws and related contextual specifications needed to govern the complex behaviour of these segments, although one can expect such a number to be quite high. Given such an abundance of segments, coupled with their (relative) freedom of occurrence and the increased number of rules, one may reasonably expect to find a certain amount of apparent matches in this corpus. The possibility of 'picking' the appropriate form to make the desired match is further increased by the existence of the forth, unspecified laryngeal – a real 'passe-partout' – and by the monosyllabic character of many of the roots (mono-syllabic roots are, in general, quite easy to match, as shown by Ringe [1999]).

The intuitive observation that, in these conditions, matches are extremely easy to find, is confirmed by a proper, quantitative

analysis. In fact, I have investigated the distribution of the laryngeal segments in **LIV**, with the purpose of quantifying the effects the use of laryngeal segments brings into the equation. The results are illustrated in the graph below, figure **V**. This graph shows that the roots displaying a higher Intrinsic Significance value (more claimed daughter language groups than variants, other things being equal; see figure **IV**), are more likely to have been reconstructed by making recourse to the laryngeal theory. This suggests that the flexibility afforded by inserting laryngeals segments into the reconstructions makes it easier to find a fortuitous match in a new language group.

Figure V: the effects of the Laryngeal Theory on the verbal roots

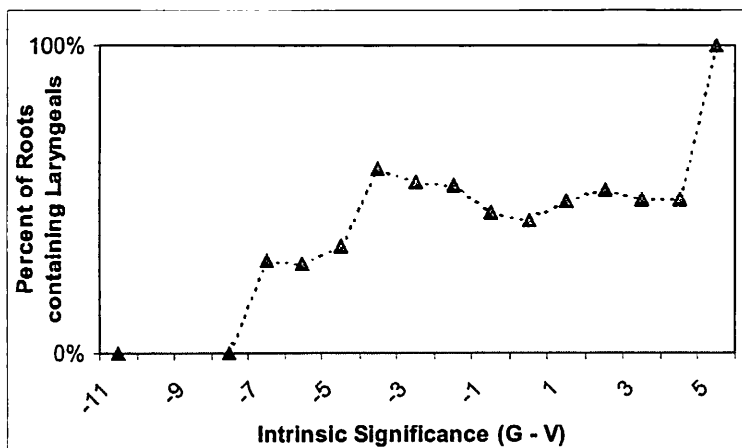


Figure V: The distribution of laryngeal roots. Those roots with higher Intrinsic Significance value are more likely to make appeal to the laryngeal theory. This suggests that the flexibility afforded by laryngeals makes it easier to find fortuitous matches.

To conclude this section it is worth observing that, despite the extra degree of flexibility granted to the overall explanatory

system, thanks to the laryngeal segments, there are still exceptions left over in the domain of vowel alternations (for which see Carruba in this volume). Sihler (1995:131) justifies the encountered exceptions as follows:

> Even though the origins of ablaut [...] were necessarily phonological, by the earliest period reachable by the comparative method the distribution of different ablaut grades in PIE had been MORPHOLOGIZED, that is, a given form or class of forms was associated with a certain grade in PIE. This is particularly true of the distribution of *e*- vs. *o*-grades [...]. As is to be expected when phonological alternations are captured by morphology, the system was never completely regular; [...]. Thus, evidence points to a non-ablauting[18] root **bhū*- or *bhuH*- 'become'

This statement, in my opinion, is a concrete testimony of how it may be very difficult, if not impossible, to falsify the IE theory, since one can always bridge the gaps between the predictions of the model and the actual data through ad-hoc explanations that will then be granted the status of a (more or less) general principle.

3. The statistically significant evidence for Indo-European

3. 1. The 'best' Indo-European roots

Let us now have a look at the most widely represented, the statistically significant roots – 6 or more branches/groups – which are listed in Table **I** below. This table shows their reconstruction, the presence *vs* absence of laryngeal segments and the various meanings of the attested forms, as given in **LIV**. In other words, this Table shows the 'Intrinsic Significance' of these roots by taking into account also all those factors that may contribute to the ease of finding a match.

At first glance these 6 roots do display a positive Intrinsic Significance value, because they are supported by a high number

of language witnesses and are reconstructed through a (relatively) small amount of variants. These are, therefore, the 'best' matches for IE. However, the high Significance value of these roots is potentially undermined by two factors: a) the frequent recourse to the laryngeals (which add an extra degree of flexibility to the explanatory system, as argued in the previous paragraph), and the high variety of meanings often involved. A root whose attested forms display a high variety of meanings clearly has less significance, because more meanings multiply the ease with which a connection can be made, as is known. The 'Significance Estimate' can be expressed through the following formula: **'G - V - M'** (Groups *minus* Variants *minus* Meanings). On the basis of this measure, for example, a simple root that meets the three-witnesses criterion and has only one meaning and one (variant) form of the root, has a significance of '1'. More meanings or variants, or laryngeal segments, reduce the Significance value because of the reasons expounded above. According to this measure then the only roots which are highly significant, being represented in all / most language groups without having to make recourse to a high number of variants are the root *$b^h er$- (which, however, has three meanings) and the root *$h_1 es$- 'to be' (which contains one laryngeal segment).

Whatever the case, and whatever relevance one might accord to factors such as presence (*vs* absence) of laryngeals and (wide) variety of meanings, the fact remains that the really statistically significant roots, according to this measure, are only just an extremely small number (only 14), with respect to the high number of roots listed in **LIV**. Last, but not least, there is to be remembered that there are often mismatches between **LIV** and other dictionaries regarding what can be considered a (more or less) regular set of correspondences, as shown, for example, by the root for 'to milk' (*$h_2 melg'$). This root is classified by Ernout-Meillet (1959) as belonging to the 'popular, familiar' variety of language and, therefore, as irregular and poorly represented across

IE. See also Buck (1949: 385-6) and Belardi (2002:140ff.) for a similar evaluation of this root, and other similar roots.

Table (I): The 'best' P-IE roots and their Intrinsic Significance value

Group G	Variant V	Meaning M	Significant G–V–M	Laryngeal	Attested meanings	page	PIE
11	3	7	1	Yes	To be: Am, is, was, etc.	214	*$h_1 es$-
7	2	4	1	Yes	To drive, lead, go, travel	227	$h_2 eg'$-
7	2	4	1		To ripen, cook, roast, bake	421	*pek^u.
8	4	2	1	Yes	To give, take	89	*deh_3.
6	4	1	1		To lick	362	*$leig'^h$
8	3	4	0	Yes	To milk, wipe away, clean, brush against	249	*$h_2 melg'$-
8	5	2	0		To sit, hold on	465	*sed-
6	4	4	0		To ask, advise, request, beg	442	*prek'-
11	9	3	-1	Yes	To know, recognize, be able to	149	*$g'neh_3$-
7	4	5	-2		To swim, float, rain, sail, flush	438	*pleu̯-
9	6	7	-4		To hit, defend, attack, wound, kill, push, scatter	195	*$gu̯^h en$-
7	6	6	-5	Yes	To grow, originate, become, be, live, sprout	83	*$b^h u̯eh_2$-
9	7	8	-6	Yes	To put, lay down, sit down, produce, make, speak, say, bring back	117	*$d^h eh_1$.
9	4	3	2		To carry, bring, take	61	*$b^h er$-

3. 2. The best Present / Perfect alternations

In Table **II** below I have listed those 8 roots which, while being statistically relevant, also appear to display what could be called a 'proper' present ~ perfect alternation. By this I mean an alternation that is not only reconstructed, not only assumed, but is actually attested in the relevant languages /language groups. Once again, at least according to the quantitative analysis carried out here, the number of roots displaying 'proper' alternations is rather low:

Table II: Present ~ Perfect alternations supported by 3 or more branches

1)**steh_2* : Pres. **sti-stéh_2* / *sth_2*- ~ *st-né* /*n̥-h_2*- vs Perf. **ste-stóh_2* / *sth_2*- ‘step (in), put (up), stand’ (alternations in Greek, Sanskrit, Latin and Celtic; p. 536)

2)**g’neh_3* : Pres. **g’n̥-né* / *n-h_3*- ~ **g’n̥h_3-sk’é* - vs Perf. **g’e-g’nóh_3* / *g’n̥h_3*- ‘recognize, know’ (alternations in Vedic, Celtic and Germanic; p. 149)

3)**telh_2* : Pres. *tl̥-né* / *n-h_2*- vs Perf. **te-tólh_2* / *tl̥h_2*- ‘pick up, take’ (alternations in Greek, Latin and Celtic; p. 565)

4)**u̯ert*: Pres. **u̯ért-e*- vs Perf. **u̯e-u̯órt* / *u̯r̥t*- ‘turn round’ (alternations in Vedic, Latin and Germanic; p. 632)

5)**b^heu̯d^h*: Pres. **b^hu-né*/ *n-d^h*- ~ **b^héu̯d^h-e*- ~ **b^hud^h-i̯é*- vs Perf. **b^he-b^hóu̯d^h* / *b^hud^h*-‘be awake’ (alternations in Vedic, Greek, Celtic and Germanic p. 66-7)

6)**b^hu̯eh_2*: Pres. **b^héu̯h_2 -e*- ~ **b^huh_2 -i̯é*- vs Perf. **b^he-b^hu̯óh_2* / **b^huh_2*- ‘grow, originate, become, be, live’ (alternations in Vedic, Latin, Greek and Germanic; p. 83)

7)**g’enh_1* : Pres. *g’i-g’énh_1/g’n̥h_1*- ~ *g’énh_1-e*- vs Perf. *g’e-gónh_1* /*g’n̥h_1*- ‘generate /produce’ (alternations in Vedic, Greek, Irish: p. 144)

8)**$h_1$ed*: Pres. **h_1ḗd* / **$h_1$éd*- vs Perf.. **h_1e-*h_1 ód* / *h_1 d*- ‘eat’ (alternations in Greek, Latin, Celtic and Germanic; p. 205)

3. 3. The nominal roots

In this last paragraph I would like to compare the state of the verbal roots with that of the nominal roots, or *lexemes* (to use a more appropriate definition), although briefly, since a detailed

analysis of this topic is outside the scope of this chapter (for a recent overview of the state of the IE lexemes see Clackson [2007:187 ff.]).

I have argued above that the great majority of the morpho-phonemic alternations reported in **LIV** are not real alternations, but rather artefacts of the method of analysis, in the sense that most variants have been ‘appended’ to the reconstruction of a given root in order to facilitate a difficult match. This thesis appears to be supported by the following fact, since long recognised in the specialist literature: ordinary nouns / lexemes display a much higher degree of irregularity than verbal roots or verbal nouns (see for example Meillet (1934: 289 ff., 379-416, 396); Benveniste (1935:175-ff.); Belardi (2002:141ff.), Campanile (1983); Ernout-Meillet (1959); Schlerath (1987) and Clackson [2007:200 ff.]). Meillet (quoted) discusses at length the ordinary nouns, including those belonging to the basic, every-day vocabulary, which he names “vocabulaire populaire”. The Author observes that most of these terms are not proper correspondences, either because they present irregularities (at least in some of the languages where they are attested), or because they cannot be safely traced back to a common, IE source (being the various terms rather different in the various branches[19]), or a combination of both factors. In addition, often these nouns only occur within three or even two contiguous languages (as already pointed out by Schmidt [1872]). These terms include the nouns for: ‘goat’, ‘dog’, ‘fox’, ‘bee’, ‘honey’, ‘milk’ (an related milking and feeding activity), ‘nail’, ‘spleen’, ‘bone’, ‘palm’ (of hand), ’sheep’, ‘night’, ‘four’, ‘wolf’, ‘bull’, ‘pig’, ‘nose’, ‘tongue’, ‘shoulder’, ‘eyes’, ‘mouth’, ‘ears’, ‘brother-in-law’, ‘horse’, as well as kinship terms referring to the semantic domain of ‘the family of the woman / wife’ and ‘distant relations’. Interestingly enough, Benveniste (1935:175), while trying to justify this state of affairs, observes that the nominal roots, unlike the verbal roots, “ne permettent pas de définir une racine, […] n’offrent pas d’alternance radicale”. By

this the Author means that the verbal roots are much more regular than ordinary nouns thanks to their typical, alternating structure. A similar way of thinking appears to lie behind the following statement by Clackson (2007:190-1):

> It is possible that [...] our rules for deriving affixed forms from roots may be a construct of the comparative process. We rely on roots as the base of derivation since we can reconstruct roots with more confidence than we can reconstruct individual lexemes. Roots may be shared across many languages, while a particular lexical formation is only found in a small number of languages

These observations, according to the analyses presented in this chapter, could be interpreted and re-phrased as follows: because of the very fact that ordinary nouns do not, indeed, display that rich alternation (supposedly) typical of the verbal roots, there is less room to 'pick and chose' among variants, and, therefore, less opportunity to justify, or 'explain away', the encountered irregularities[20]. Not only, the fact that it is much more difficult to reconstruct "individual lexemes" may be an indication that the IE roots, as conventionally reconstructed (including their assumed, intricate and often unjustified derivational procedures) are indeed an artefact of the method of analysis.

4. Conclusion

The quantitative analyses carried out above show that the lexical comparative corpus and the present / perfect, *e* ~ *o* alternation assembled in **LIV** definitively lack statistical significance. It is reasonable therefore to assume that most of the correspondences and most of the present ~ perfect alternations under discussion are chance resemblances, artefacts of the traditional method of analysis. However, these very same analyses have also individuated those roots and those alternations that do display

statistical significance and are therefore most likely to represent genuine linguistic correlations, although the number of these 'good' roots and alternations is lamentably low.

What are then the conclusions to be drawn from this? The conclusions to be drawn may vary according to which position scholars hold with regard to the following, basic methodological issues. How relevant, how 'diagnostic' are the lexical correspondences in general and therefore within IE (as against the morphological correspondences)? How many good correspondences, if any, are required, or at least desirable, in order to establish and 'prove beyond reasonable doubt' a language family? How relevant are the morpho-phonological correlations of the type examined above for the task of assessing genetic relations? And, if they do rate quite high in this task, would that small bunch of present / perfect, 'proper' alternations cutting across several (but not all) IE languages constitute a diagnostic clue of relatedness? Is the comparative method a heuristic method or not? Are 'obvious', 'compelling' correlations good enough to assume, if not prove beyond doubts, genetic relatedness? Etc. If no consensus is attained on these fundamental issues, drawing 'a' specific conclusion from the analyses carried out here will turn out to be impossible.

Whatever the answers to these questions may be, my personal opinion is that the circularity issue embedded in the comparative method has not (yet) been resolved and that, as a consequence, the IE theory, as it stands today, is unable to make clear-cut and testable predictions. On the contrary, the IE theory appears to be still flexible enough to be adjustable (and adjusted) to account for almost any data. To use the words of an eminent physicist, W. Pauli[21], a theory of this sort may be regarded as: "not even wrong".

Endnotes

[1] There is no room here to give a comprehensive account of the long-standing debate regarding the scientific nature, or otherwise, of the comparative method; see *Introduction* in this volume, Koerner (1989) and Salmons & Joseph (eds, 1998).

[2] As Fox (1995: 63) puts it: "An inevitable problem is that of circularity: it is difficult to avoid some version of the vicious circle that results from assuming that forms are cognate because they can be reconstructed with the same proto-phoneme, where the proto-phoneme is itself the result of assuming that they *are* cognate".

[3] See *Introduction* for a fuller version of this quote

[4] For a detailed account of the history of quantitative, historical linguistics – scope of applicability, methods and results – see McMahon & McMahon (2003) and (2005); see also the chapter by Drinka in this volume.

[5] As far as I know, the only attempts to assess the statistical significance of an IE corpus have been made by Bird (1982) and then by Ringe (1995), although the purpose of their analysis is not that of verifying the soundness of the IE theory. The purpose of these studies is that of comparing the distribution of cognates in Nostratic with the distribution of cognates in "uncontroversial language families" (Ringe 1995:64), such as IE, in order to show the difference between the two cases. More specifically, Ringe analyses the comparative corpus of Pokorny (1959-69), using the statistical summary of the distribution of cognates in the dictionary as carried out by Bird (1982). The purpose it to investigate whether the distribution of cognates in Pokorny does produce a curve that can be generated randomly, that is, whether these cognates might be (at least in part) the effect of chance resemblances, and then to compare the IE situation with the distribution of cognates within Nostratic. Ringe (1995:65) concludes that: "Pokorny's evidential standards are lamentably lax; he includes a large number of items which probably should *not* be reconstructed for Proto-Indo-European (PIE), either because the cognate set shows too many irregularities of one sort or another, or because 'sound symbolism' of various kinds can account for the similarities observed.

[6] Note that, strictly speaking, **LIV** does not list 'indogermanischen Wurzeln', but 'indogermanischen Verben', that is, those roots provided with 'Primärstämme' (see below). Note also that I have not used the second edition of **LIV** (2001) because it was not yet available when this research was carried out.

[7] Hock (1993) presents a long list of what he considers to be chance resemblances between Tamil and English as well as Indo-Aryan and English. Hock (1994) presents a list of Dravidian /Indo-Aryan chance similarities. He shows how difficult it can be to tell apart proper correspondences, borrowing and chance resemblances not only within macro-families but also within well established families.

[8] The fact that many alternating starting-point reconstructions can be too a powerful tool of analysis has also been pointed out within Uralic studies, where most scholars refrain from reconstructing proto-alternations, that is, attributing them to the Uralic proto-language.

[9] The question mark is in the text, meaning that the reconstruction is considered by **LIV** as unsure.

[10] There are 19 different types of present classified in **LIV**.

[11] The roots that (supposedly) tend to display internal alternations are the so-called a-thematic ones, that is, those built without thematic vowel. They are characterized by having 'strong' forms, that is, full grade, at the singular, and 'weak' forms elsewhere (including passive and reflexive).

[12] It is believed that the vocalism *-a-* in the singular represents indeed an original *-ŏ-*, and not an original *-ĕ-*, because this is what is (indirectly) demonstrated by the absence of palatalization of the initial, radical velar, as in *cakára*.

[13] Greek documents directly a radical grade *-ŏ-* for the three singular persons, although only in a few verbs, such as the form mentioned in the text and other, Homeric verbs: ἔοικα : ἔικτον, μέμονα : μέμαμεν, etc. In many other verbs the grade *-ŏ-* has been generalized to the entire declension.

[14] Like Latin, Old Greek has actually developed several types of different, secondary perfects.

[15] The situation regarding the 'strong preterits' is more complex; however, there appears to be consensus on the fact that this tense cannot be traced back directly to the IE perfect; see for example Di Giovine (1996:139 ff.) and the bibliography quoted there.

[16]"In the conjugation of the Latin Perfect one cannot find safe examples of the continuation of an **-ŏ-* grade in the radical syllable… and there are no examples of alternation between strong and weak forms within the conjugation of the *perfectum*'.

[17]In translating the meaning of the various forms, I have followed **LIV** way of indicating –at times – also the (I, II, III) person together with the meaning of a given root (although **LIV** itself is not always consistent).

[18] See the reconstruction of the root **b^hu̯eh₂-* according to **LIV,** given in Table (**I**). See also the comment by Clackson (2007: 87), according to whom certain roots (as well as nominal and verbal suffixes and endings) which theoretically should show ablaut forms in fact never do, such as, indeed, the non-ablauting form **b^huH-*, present in all IE branches except Anatolian and Tocharian.

[19] This diversity, of course, can also be the effect of added, different affixes to an original, common root and this (it has been claimed) can be the origin of the variety of meanings. This may be true, although it does not help in the statistical analysis.

[20] Even if one does not accept the explanation proposed here, the fact remains that the basic IE lexicon is tainted with variation, instability and irregularity, and

this is in stark contrast with the picture usually presented in text books, where the IE basic terms are considered to be examples of perfect correspondences. Textbooks generally report quite a short list of basic terms, selecting them among the best sets of correspondences available, whilst omitting to mention that most of them display (more or less serious) difficulties. An example of this is the Table titled: "Some Basic Indo-European Terms" by Baldi (1987: 24-25), that includes the terms for 'dog', 'horse', 'wolf', 'sheep', pig', 'tongue' and 'four'. Some of these terms contain irregular sound changes that, admittedly, have defied any attempt to justification, such as the terms for 'wolf', from proto-Germanic **wulfaz* < P-IE **wḷkʷos* (see for example Sihler [1995:157]).

[21] The statement: "not even wrong", attributed to W. Pauli (according to Rees [1999/2000:76]), does not refer, of course, to the IE theory, but to a specific theory within physics.

References

Baldi, P.
1987 Indo-European Languages. In Comrie, B. (ed.), *The Major Languages of Western Europe*. London: Routledge. 21-58.

Belardi, W.
2002 *L'Etimologia nella Storia della Cultura Occidentale*, I. Roma: Il Calamo.

Bird, N.
1982 *The Distribution of Indo-European Root Morphemes*. Wiesbaden: Harrassowitz.

Benveniste, E.
1935 *Origines de la Formation des Noms en Indo-Européen*. Paris: Librairie Adrien-Maisonneuve.

Brady, R. & Marcantonio, A.
2003 Evidence that most Indo-European reconstructions are artefacts of the linguistic method of analysis. In Hajičová, E. et al. (eds), *Proceedings of the 17th International Congress of Linguists*. Prague: Matfyzpress, MFF UK. CD-ROM ISBN: 80-86732-21-5.

Buck, C. D.
1949 *A Dictionary of Selected Synonyms in the Principal Indo-European Languages*. The University of Chicago Press.

Chrétien, C. D.
1937 Quantitative classification of Indo-European languages. *Language* 13:83-103.

Campanile, E.
1983 *Problemi di Lingua e di Cultura nel Campo Indoeruropeo*. Pisa: Giardini.

Campbell, L.
1998 *Historical Linguisitcs. An Introduction*. Edinburg University Press.

Clackson, J.
2007 *Indo-European Linguistics. An Introduction*. CUP.

Collinge, N. E.
1985 *The Laws of Indo-European*. Amsterdam: Benjamins.

Di Giovine, P.
1996 *Studio sul Perfetto Indo-Europeo*, II. Roma: Il Calamo.

Ernout, E. & Meillet, A.
1959 *Dictionnaire Etymologique de la Langue Latine*. Paris: Klincksieck (4th edition).

Fox, A.
1995 *Linguistic Reconstruction. An Introduction to the Theory and Method*. Oxford University Press.

Grace, G.
1990 The "aberrant" versus "exemplary" Melanesian languages. In P. Baldi (ed.), *Linguistic Change and Reconstruction Methodology*. Berlin: Mouton de Gruyter. 155-173.
1996 Regularity of change in what? In M. Ross & M. Durie (eds), 157-179.

Greenberg, J. H.
2000 *Indo-European and its Closest relatives: The Eurasiatic Language Family*. Vol. 1: *Grammar*. Stanford University Press.
2001 *Indo-European and its Closest relatives: The Eurasiatic Language Family*. Vol. 2: *Lexicon*. Stanford University Press.
2005 *Genetic Linguistics: Essays on Theory and Method*. Edited by W. Croft.

Harrison, S. P.
2003 On the limits of the comparative method. In B. D. Joseph & R. D. Janda (eds), *The Handbook of Historical Linguistics*. Oxford: Blackwell. 213-243.

Hock, H. H.
1993 SWALLOW TALES: Chance and the "world etymology" MALIQ'A 'swallow, throat'. In K. Beals et al. (eds), *Papers from the 29th Regional meeting of the Chicago Linguistic Society*. 215-238.
1994 (Pre-)Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. *Studies in the Linguistic Sciences* 14/1: 89-107.

1996 Pre-ṛigvedic convergence of Indo-Aryan and Dravidian? A survey of the issues and controversies. In J. E. M. Houben (ed.), *Ideology and Status of Sanskrit*. Leiden: Brill. 17-58.

Illič-Svityč, V. M.

1971-1984 *Opyt sravnenija nostratičeskikh jazykov (semitokhamitskij, kartvelskij, indoevropejskij, ural'skij, dravidskij, altajskij)*, I-III. Moscow: Nauka.

Koerner, E. F. K.

1989 Comments on reconstruction in historical linguistics. In Vennemann, T. (ed.), *The New Sound of Indo-European: Essays in Phonological Reconstruction*. Berlin / New York: Mouton de Gruyter. 1-20.

LIV = Rix, H. (ed.)

1998 *Lexikon der Indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen*. Wiesbaden: L. Reichert.

Macdonell, A. A.

1995 *A Vedic Grammar for Students*. Delhi: Motilal Banarsidass.

Marcantonio, A.

2002 The Uralic Language Family: Facts, Myths and Statistics. *Transactions of the Philological Society*. Oxford/Boston: Blackwell.

2003/2005 Evidence: The missing concept in comparative studies: A preliminary comparison of Uralic and Indo-European. In M. M. J. Fernandez-Vest (ed.), *Les Langues Ouraliennes Aujourd'hui: Approche Linguistique et Cognitive / The Uralic Languages Today: a Linguistic and Cognitive Approach*. Bibliothèque de l'École des Hautes Études: Sciences Historiques et Philologiques 340. Paris: Librairie Honoré Champion. 117-132.

Masica, C. P.

1979 *Defining a Linguistic Area: South Asia*. Chicago University Press.

McMahon, A. & McMahon, R.

2003 Finding families: quantitative methods in language classification. *Transactions of the Philological Society* 101/1: 7-57.

2005 *Language classification by Numbers*. OUP.

Meillet, A.

1934 *Introduction à l'Étude Comparative des Langues Indo-Européennes*. Paris: Librairie Hachette.

Morpurgo Davies, A.

1978 Analogy, segmentation and early Neogrammarians. *Transactions of the Philological Society. Commemorative volume: The Neogrammarians*. 36-64.

Nichols, J.

1996 The comparative method as heuristic. In Ross, M. & Durie, M. (eds), 1996. *The Comparative Method Reviewed. Regularity and Irregularity in Langue Change*. OUP.

Pokorny, J.

1959-69 *Indogermanisches etymologisches Wörterbuch*. Bern: Francke.

Prokosch, E.

1939 *A comparative Germanic Grammar*. Philadelphia: University of Pennsylvania Press.

Rees, M.

1999/2000 *Just six numbers*. London: Phoenix.

Ringe, D.

1992 On Calculating the Factor of Chance in Language Comparison. *Transactions of the American Philosophical Society* 82: 1-110. Philadelphia.

1993 A reply to Professor Greenberg. *Proceedings of the American Philosophical Society* 137/1: 91-109.

1995 "Nostratic" and the factor of chance. *Diachronica* 12 / 1: 55-74.

1998 A probabilistic evaluation of Indo-Uralic. In J. C. Salmons and B. D. Joseph (eds), 153-198.

1999 How hard is it to match CVC-roots? *Transactions of the Philological Society* 97:213-244.

Ringe, D., Warnow, T. & Taylor, A.

2002 Indo-European and computational cladisitcs. *Transactions of the Philological Society* 100/1: 59-129.

Salmons, J. C. and Joseph, B. D. (eds)

1998 Nostratic: Sifting the Evidence. *Amsterdam Sudies in the Theory and History of Linguistic Science. Series IV. Current Issues in Linguistic Theory*. Amsterdam: J. Benjamins.

Schmidt, J.

1872 *Die Verwandtschaftsverhältnisse der indogermanischen Sprachen*. Weimar: H. Böhlau.

Schlerath, B.

1987 On the reality and status of a reconstructed language. **JIES** 15: 41-6.

Sihler, A. S.

1995 *New Comparative Grammar of Greek and Latin*. OUP.

Szemerényi, O.

1996 *Introduction to Indo-European Linguistics*. OUP.

Unger, J. M.
1990 Summary report of the Altaic panel. In P. Baldi (ed.), *Linguistic Change and Reconstruction Methodology*. Berlin: Mouton de Gruyter.

Defining the Limits of Grammatical Borrowing

Yaron Matras
University of Manchester

An examination of Romani and Domari, two Indo-Aryan languages spoken outside of India, shows that borrowing has an impact on a wide range of grammatical domains and categories. This includes even the domain of bound, inflectional morphology, often thought of as relatively immune to contact influence. Nonetheless, borrowing is not random, but tends to be structured in a hierarchical manner and so it is at least to some extent predictable. This suggests that, to the historical-comparative linguist, some components of grammar offer more reliable indicators than others about shared historical-genetic inheritance.

1. Introduction

This paper attempts to contextualise contact-induced change in the domain of grammar – 'grammatical borrowing' – by examining two case studies: The European dialects of Romani, and Domari, both territorially isolated New Indo-Aryan languages. These languages stretch both grammatical and lexical borrowing almost to their very limits, yet there are boundaries to the extent of borrowing even in these cases. In particular, it will be shown that contact influences in grammar do not occur at random, but tend to follow rather hierarchical patterns that are to some extent at least predictable. Given the general orientation of the volume toward aspects of historical reconstruction and the less contested observation that clause-level syntax is a weak indicator of genetic relatedness, my focus will be on morphology and grammatical (morpho-) lexicon.

2. The languages

Both Romani and Domari are descendents of Central Indo-Aryan languages dating back to the transition period from Old Indo-Aryan (OIA) to Middle Indo-Aryan (MIA). Both languages share their most ancient layer of innovations with Central Indo-Aryan: Syllabic *ṛ* in OIA *śṛn-* 'to hear' develops in both languages into a raised vowel (Romani *šun-*, Domari *sin-*). The cluster in OIA *akṣi* 'eye'is simplified to *k* (Romani *jakh*, Domari *iki*), and the cluster in OIA *asmnan, tusme* 'we, you.PL' loses its fricative segment (Romani *amen*, Domari *eme* 'we', Romani *tumen*, Domari *itme* 'you.PL'). The initial semi-vowel in OIA *yuvatiḥ* 'woman' becomes an affricate (Romani *džuvel*, Domari *džuwir*). This combination of isoglosses places the ancestral forms of both languages in close proximity to those of the present-day Central languages of India (cf. Masica 1991).

Both languages also preserve a number of archaic OIA features, such as the clusters *st* in Romani *vast*, Domari *xast* 'hand', and *dr* in Romani *drakh*, Domari *drak* 'grape', as well as intervocalic dentals such as Romani *gelo*, Domari *gara* 'gone' (OIA *gata*), all features that are not retained in the modern Central languages. Both languages also preserve, by and large, the consonantal present-tense person concord set, as in (1):

(1) Present-tense person concord (analogous formations and innovations are shaded)

	OIA	Romani	Domari
1SG	*-ami*	*-av(a)*	*-ami*
2SG	*-asi*	*-es(a)*	*-ek*
3SG	*-ati*	*-el(a)*	*-ari*
1PL	*-amas*	*-as(a)*	*-ani*
2PL	*-ati*	*-en(a)*	*-asi*
3PL	*-anti*	*-en(a)*	*-andi*

Alongside these conservative traits, which indicate a rather early migration away from the Central regions, both languages participate in other changes that are typical of the transition period from MIA to New Indo-Aryan (NIA), such as reduction of gender and case marking in morphology, and the simplification of consonant clusters, in phonology (OIA *sarpa* 'snake', MIA *sappa*, Romani and Domari *sap*, cf. Hindi *sāp*).

Two rather diagnostic innovations are shared, to some extent at least, with languages of the extreme Northwest of India (e.g. Kashmiri), often referred to as the Indo-Iranian 'frontier languages' or 'Dardic' (cf. Grierson 1922). The first is the progression of grammaticalisation of postposed case particles into synthetic case markers, which in both languages is well-pronounced and has resulted in the emergence of not just one or two, but a whole series of five agglutinating suffixes (2):

(2) Romani and Domari Layer II case affixes

Case	Romani	Domari
Benefactive	*-ke/-ge*	*-ke*
Dative	*-te/-de*	*-ta*
Locative		*-ma*
Sociative	*-sa/-ca*	*-san*
Ablative	*-tar/-dar*	*-ki*
Genitive	*-ker-/-ger-*	

The second development is even more diagnostic of a close affinity to the 'Dardic' or North-western frontier languages in their early modern period. It concerns the emergence, post-ergativity, of a new past-tense conjugation, based on the attachment of oblique pronominal clitics to the past participle of the verb. The development is only partly complete, to different extents in each language (and with some variation within Romani). Typically, in

the third person the past tense of verbs of motion and change of state continues the old formation, in which the participle appears on its own, in an adjectival form, which agrees with the subject/undergoer: Romani *gel-o*, Domari *gar-a* 'he went', Romani *gel-i*, Domari *gar-i* 'she went'. In transitive verbs, and in other persons also in all other verbs, the subject is represented by an historical oblique pronominal clitic, which has merged with the participle to form a new set of person concord markers (3):[1]

(3) Romani and Domari past-tense person concord and its historical predecessors (for the verb 'to do') (analogies and innovations are shaded)

Person	MIA participle	MIA oblique pronoun	Romani	Domari
1SG	*ka(r)da*	*me*	*kerdjom*	*kardom*
2SG		*te*	*kerdjal/-an*	*kardor*
3SG		*se*	*kerdjas*	*kardos*
1PL	*ka(r)de*	*ṇe*	*kerdjam*	*kerdēn*
2PL		*ve*	*kerdjan*	*kerdēn*
3PL			*kerde*	*karde*

The trail of archaisms, innovations and related isoglosses leads us to trace a very similar pathway in the development of the ancestral idioms of both Romani and Domari: Both seem to have broken away from an ancient Central Indo-Aryan cluster of dialects during the early transition period to MIA, but remained within reach of pan-Indo-Aryan innovations as late as the late MIA period and even the emergence of early New Indo-Aryan. While some of the conservatisms are idiosyncratic in the two languages, others are shared with languages of the Northwest, with which, in addition, some diagnostic late innovations are also shared.

In light of this very particular reconstructed history it is tempting to assume that Romani and Domari were both part of the very same lineage, and that they parted company only at a late stage in their development, possibly after leaving India. This was indeed the view taken by Sampson (1923) and other contemporaries. Turner (1926) however pointed out differences between the languages in the treatment of retroflex sounds (cf. Romani *ařo*, Domari *ata* 'flour', from *āṭa*; Romani *phuřo*, Domari *wida* 'old man', from *buḍḍha*), arguing for a rather old differentiation. Hancock (1995) has, in addition, pointed out that the inventory of Iranian loanwords in Romani and Domari is quite distinct, suggesting a split before Iranian influence and so before the exodus from Indian-speaking territories.

Outstanding differences also exist in grammatical vocabulary as well as in the position of pronominal affixes (which in Domari attach to nominal entities and prepositions, as well as to verbs). Other differences include the apparently rather early gramamticalisation in Romani of *-(j)ov-* (from OIA *bhuv-* 'become') as a mediopassive affix, while Domari retains OIA *-y-* (*bag-* 'to break', *bagy-* 'to be broken'); the animacy constraint on the use of direct object marking in Romani (which is missing from Domari); the grammaticalisation of the verb *čh-* 'to stay' to a subjunctive marker in Domari intransitive verbs; the formation of the nominative plural in nouns (Romani *-a*, Domari *-e*); early reflexes of OIA *v-* and of OIA sibilants such as *ṣ* and *ś* (Romani *berš*, Domari *wars* 'year', OIA *variṣa*), and more (cf. Matras 1999).

The most likely historical scenario therefore seems to be an origin in a shared socio-ethnic rather than strictly linguistic community, with a shared migration history within India leading to participation in the same set of isoglosses, followed by a similar pattern of westwards migration out of India and subsequent retention of the original language under diverse foreign influences. In the socio-historical context of the Indian sub-continent, such a shared social identity that does not necessarily imply linguistic

unity can be fairly easily traced to the caste system with its numerous sub-divisions. This is confirmed by the employment in both communities of self-appellations deriving from the Indian caste-name *ḍom* (Romani *řom*, Domari *dom*), and moreover by the presence of a wholesale label for outsiders who are no caste-members, Romani *gadžo*, Domari *kaddža*, which has cognate expressions in the languages of present-day communities belonging to the *ḍom* and similar caste groups in India. The presence outside of India of other groups of Indian origin that carry caste-names as self-appellations, such as *Lom* in the Caucasus, *Jat* in Afghanistan or *Parya* in Tajikistan, confirms an overall phenomenon of emigration from India of specific caste groups and the maintenance of caste-like identity even after the breakaway from the actual caste-based social system of the Indian subcontinent.

In their subsequent history the two languages differ considerably. Earlier forms of Romani absorbed influences from Iranian and Armenian, as well as a very strong Byzantine Greek element, including many elements of vocabulary and even inflectional material. From the early fourteenth century onwards, Romani-speaking populations spread throughout Europe, leading to the dispersion and subsequent differentiation of the language into very distinct dialects. While Romani spread to virtually all regions of Europe, a high population density remained in central-eastern and South-eastern Europe. With well over 3.5 million speakers and possibly many more, Romani is now the largest minority language in the European Union. It is gradually occupying a space in the public domain, with numerous initiatives promoting literacy in the form of publications, websites, broadcasting and other media, as well as initiatives to introduce the language into the curriculum. Codification is generally region-based, with no uniform standard. Romani has been the subject of intense research dating back to the early eighteenth century, with an upsurge of interest in the modern period from 1990 onwards.[2]

Domari, by contrast, is documented primarily in work on the variety spoken in Jerusalem (Macalister 1914, Matras 1999). Fragmented documentation of word lists and individual phrases allows us to reconstruct a picture of a spread of the language, probably in different varieties, from the Caucasus in the north and as far as Sudan in the south. It is reported that the language is still being spoken in communities in eastern Anatolia (Turkey), Lebanon and Syria, and Jordan. In most communities, it seems that use of the language is limited to the older generation. Jerusalem Domari is certainly highly endangered and is only spoken by a very small circle of elderly members of the community; the dominant family language is now Arabic. From Jerusalem Domari and other fragmented documentation of Palestinian Domari it appears that the language had been in close and prolonged contact with southern Kurdish, before immigration into Arabic-speaking territory. A layer of Turkish loanwords testify in all likelihood to the trade activities of the Dom as nomadic smiths and musicians under the Ottoman rule, but possibly also to close social links with nomadic Turkmen.

The following sections survey the contact behaviour of Romani and Domari in respect of selected functional categories. The comparison is biased by the existence of a large corpus of descriptions of Romani dialects, including the RMS (Romani Morpho-Syntax) Database with information on over 100 varieties of the language, which tags contact influences systematically (see Elšík & Matras 2006), while for Domari the material covers only the variety of a single location. Moreover, given the highly differentiated contact behaviour of Romani varieties in some domains, it is certainly possible that Jerusalem Domari may not be representative of Domari as a whole. Nonetheless, the enormous impact of contact both on Romani as a whole and on Jerusalem Domari, and the similar time-depth, socio-economic history, and ultimate genetic-typological origin of both languages justify a comparative discussion even in the absence of equal or even near-equal dialectological coverage.

3. Inflectional morphology

Inflectional morphology is well-known to be relatively resilient to borrowing, and therefore a rather stable indicator of genetic inheritance, though naturally it is subject to language-internal renewal and so it poses other challenges in the way of historical reconstruction. I therefore begin my survey of grammatical borrowing with this domain. The resistance of inflectional morphology to borrowing is best confirmed in the two languages in the area of case inflection: in neither of the two languages is there any direct attestation of borrowing of case markers. The closest to case marker borrowing is the apparent extension of what is a strongly lexicalised, inherited Romani marker of direction and location *-e* (e.g. *kher* 'house', vs. *kher-e* 'home, at home'), in Zargari, a Balkan dialect of Romani that has migrated eastwards and is now spoken in Iranian Azerbaijan (Windfuhr 1970, Baghbidi 2003). Here, it is used much like a dative or directional case, possibly by analogy to the Azeri case marker *-e/-a*. The potential for the borrowing of case-markers is of course limited due to the paucity of transparent case markers in many of the contact languages. Those European languages that do have case inflection tend to possess highly flectional segments whose semantic transparency is rather low, a factor which inhibits borrowings (cf. Field 2002). Nonetheless, no borrowing is attested from languages like Finnish, Hungarian, Basque or Turkish either.

Borrowed nominal morphology is limited to nominative singular and plural endings. Domari incorporates Arabic nouns along with their plural ending, but adds its own plural formation to that: *zálame* 'man', plural *zlām-é* (from Arabic *zálame* 'man', plural *zlām*, with addition of the Domari plural ending *-e*). Romani shows productive use of borrowed nominal inflection endings. Early (Medieval) Romani borrowed Greek nouns along with their nominative inflection markers, thus *for-os* 'town' (from the Greek word for 'market'), plural *for-i*, *kokal-o* 'bone', pl. *kokal-a*. These

Greek-derived inflectional endings continue to be productive in the European dialects of Romani and serve as the basis for the integration of further loanwords from contemporary contact languages: e.g. *prezident-os* 'president', *doktor-is* 'doctor', *šeft-o* 'deal' (German *Geschäft* 'business'). Some individual Romani dialects continue to enrich the inventory of plural markers through borrowings from the contemporary contact languages. Vlax Romani adopts the Romanian-derived plural suffix *-uri/-ura*, which is employed both with earlier loans, e.g. *foruri* 'towns', and with new loans, as in *šefturi* 'deals' (but not combine with pre-European vocabulary). Epirus Romani, by contrast, adopts the Greek plural ending *-imata* which is applied to all inherited masculine nouns ending in a consonant: *vast* 'hand', PL *vastimata*; *kher* 'house', PL *kherimata*. Kaspichan Romani (northern Bulgaria) uses Greek-derived singular and plural endings with loans from Turkish: *džam-is* 'mosque' (Turkish *cami*), PL *džam-ides*.

Romani also shows borrowing in the domain of verb inflection. The Greek tense-aspect inflectional markers *-Vn-*, *-Vz-* etc. (present) and *-is-* (past) entered the language in all likelihood along with Greek-derived verbs, and were subsequently generalised to verbs from other contact languages. We find forms like *analadi-s-ker-djom* 'I understood', where the form *anladi* is the Turkish inflected past-tense 3SG form *anal-dı* 'understood', the *-(i)s-* is the Greek aorist marker, *-ker-* is the inherited causative/transitive marker that integrates the loan verb, and *-djom* the inherited past-tense 1SG. These Greek markers have thus assumed a role within Romani morphology that is partly inflectional (indicating tense) and partly derivational (indicating particular verbs as loans).

Borrowed person concord markers on the verb appear marginally as well. Greek has contributed the 3SG present-tense marker *-i*. In many Romani dialects it is limited to loan verbs, e.g. Arli Romani of Kosovo *pomožin-i* 'he/she helps'. In Slovene Romani it is generalised to all verbs, replacing the inherited Romani 3SG present-tense concord marker.

Slovene Romani also borrows the Slovene/Croatian person concord endings for the 2PL – *kerdž-ate* 'you.PL did', and the 1PL – *mothav-amo* 'we say'. A cluster of Romani dialects in contact with Turkish, mainly in northern Bulgaria, borrow elements of their person concord set from Turkish, resulting in hybrid past-tense formations such as *kerd-aməs* 'we did' (Romani *kerd-am*, Turkish *yapt-ımız*) and *kerd-enəs* 'you.PL did' (Romani *kerd-en*, Turkish *yapt-ınız*). This process is triggered by analogies between the Romani 1st and 2nd person forms in *-m* and *-n*, respectively, and the corresponding Turkish forms in *-m* and *-n*, which is no doubt supported by the presence of similar nominal possession suffixes in Turkish. This analogy triggers a re-interpretation of the Turkish form *-Vz* as a generic plural marker, which is then attached to the inherited Romani forms (4):

(4) Borrowing of Turkish person concord markers into Romani

	Romani (inherited)		Turkish			Romani new form
Person	Present	Past	Present	Past	Nominal	Past
1SG		*-om/-em/-im*	*-Vm*	*-Vm*	*-Vm*	
2SG		*-an*	*-sVn*	*-Vn*	*-Vn*	
1PL	*-as*	*-am*	*-Vz*		*-VmVz*	*-am-əs*
2PL	*-en*	*-en*	*-sVnVz*	*-VnVz*	*-VnVz*	*-en-əs*

Even more widespread in the Balkans is the wholesale retention of Turkish verb inflection with lexical verbs borrowed from Turkish, resulting in the compartmentalisation of the verb inflectional domain into two separate sets of conjugations, inherited and borrowed. This can be illustrated by the contrast between the following two verbs in the Kalburdžu Romani dialect of Sindel, Northeastern Bulgaria (5):

(5) Inherited and Turkish-derived verb conjugations in the Kalburdžu Romani dialect of Sindel, Northeastern Bulgaria

	Inherited: *phurjo(v)-* 'to grow old'		Turkish: *evlen-mek* 'to marry'	
Person	Present	Past	Present	Past
1SG	*phurjovav*	*phurilem*	*evleniim*	*evlendim*
2SG	*phurjo*	*phurilan*	*evlenisin*	*evlendin*
3SG	*phurjol*	*phurila*	*evlenii*	*evlendi*
1PL	*phurjova*	*phurilam*	*evleniis*	*evlendik*
2PL	*phurjon*	*phurilen*	*evlenisinis*	*evlendinis*
3PL	*phurjon*	*phurile*	*evleniler*	*evlendiler*

The incipient use of original verb inflection with loan verbs can be observed in Russian Romani as well as in the Romani dialects of Greece. In the Greek Romani dialect of Epirus, for example, some Greek-derived verbs are adapted to Romani by means of loan verb adaptation markers – deriving, ironically, from Greek tense/aspect markers – but take inherited Romani inflection – *parakal-iz-ava* 'I thank' (Greek *parakal-o*), *aɣap-ez-ava* 'I love' (*aɣap-o*) – while others are inserted along with their Greek inflection:

(6)

na *bor-o* te *diavaz-o* soske *prep-i*
NEG can-1SG COMP study-1SG because must-3SG

te vojt-iz-av me daj-a
COMP help-LOAN-1SG my-OBL mother-OBL

I *cannot study* because I *have* to help my mother'

The modal verbs *boro* 'I can' and *prepi* 'must' (impersonal) in (6) take Greek inflection, as does the lexical verb *diavazo* 'I study'; but the Greek-derived lexical verb *vojt-* 'to help' is integrated into inherited Romani inflection.

The use of original, source-language inflection with borrowed modals is much more common than with borrowed lexical verbs, and is found in a number of Romani dialects. Gurbet Romani in Serbia, for example, replicates alongside the Serbian 3SG form *mora* 'must' also the Serbian 1SG *moram* 'I must' and 2SG *moraš* 'you must'. In Domari, the borrowed Arabic modals *lāzim* 'must' and *mumkin* 'it is possible' are impersonal. But other modals and auxiliaries borrowed from Arabic take full Arabic person and tense inflection. They include the aspectual auxiliary *kān* 'was', which indicates past-habitual (*kān džari* 'he used to go', *kānat džari* 'she used to go', etc.), the transition-modals *ṣār* 'to begin' and *baqi* 'to continue', as well as the nominal form *bidd-ī* 'I want', *bidd-ak* 'you.M want', etc. It is noteworthy that the incorporation of Arabic inflection introduces a gender distinction into Domari in positions in which it is otherwise not indicated in the language (i.e. with present-tense verbs). The result is a compartmentalisation of the system of verbal morphology, separating most modal and auxiliary verbs, which are borrowed from Arabic and take Arabic inflection, from all lexical verbs (including those that are borrowed from Arabic), which take the inherited (Indic) inflection.

4. Deictic and anaphoric expressions

Deictic and anaphoric forms are rare among confirmed borrowings. Some Romani dialects of the Balkans have the place deixis *orde* 'there', possibly from Turkish or Azeri *orda/ordä*, as well as *inća* 'here', possibly from Persian *īndžā*, though both etymologies require further investigation. Some varieties of Sinti Romani borrow German-derived *doti* 'there'. Rumungro (Selice)

Romani borrows the Hungarian deictic prefixes *am-* and *ugyan-*, which are combined with Romani deictic stems (Elšík 2007). Domari makes occasional use of the Arabic resumptive pronoun *iyyā-*, complete with its Arabic agreement inflection, in relative clauses. Domari also employs Arabic the reflexive expression *ḥāl-* '-self' and the Arabic reciprocal *baʕḍ*.

A number of Romani dialects show the development of hybrid pronominal forms (7):

(7) Borrowed pronominal forms in Romani and their origin:

	3SG	3PL	Source	3SG	3PL
Early Romani	*ov*	*on*			
Hungarian Romani	*ov*	*on-k*	Hungarian	*ő*	*ő-k*
Slovene Romani	*ov*	*on-i*	Slovene	*on*	*on-i*
Thracian Romani	*ov*	*on-nar*	Turkish	*o[n-]*	*on-lar*
Molise Romani	*jov*	*lor*	Italian	*il*	*loro*

Early Romani (see Matras 2002, Elšík & Matras 2006) is assumed to have had the pronominal forms *ov* 'he', *on* 'they', which are continued in most dialects of South-eastern Europe, while in other areas prothetic consonants *j-* and *v-* are added. The remarkable development attested in the top three dialects – Hungarian, Slovene, and Thracian Romani – is the copying of plural affixes from the contact languages into the inherited plural pronominal form. This is prompted by the accidental similarity of forms between the Romani pronouns and those of all three contact languages, as well as by the presence of an agglutinating formation

of the plural pronoun in the contact language, which is identical to the general, nominal plural suffix used in the language. It is thus the plural affix, not the actual pronominal form of the respective contact language that is borrowed into Romani. The exception is the borrowing of a word-form in Molise Romani, which is the only genuine case of the borrowing of a pronominal word-form. Molise Romani also borrows the reciprocal and 1PL reflexive form *ci* from Italian.

5. Derivational morphology

Romani dialects borrow a series of agentive and diminutive affixes from various contact languages. Common in particular are diminutive/diminutive feminine *–ic--ica/-icka-* and agentives *-ari* and Turkish-derived *-dži*, which are shared by several languages in the Balkans. Borrowing of other nominal derivation is less common. Borrowing of verb-deriving morphology is quite rare. A rather isolated example is Lovari Romani *-áz-* from Hungarian, found in a limited number of inherited words, such as *bučáz(in)-* 'to work', from *buči* 'work'.

Romani dialects in contact with Slavic languages primarily in Poland, Russia, and Slovakia, and to some extent also Romani dialects in contact with Latvian, Lithuanian, and Greek, borrow *Aktionsart*-derivational prefixes (often referred to as 'aspect'): In Russian Romani we find, based on *dava* 'I give', the verb *dodava* 'I add', *obdava* 'I embrace', *otdava* 'I confiscate', *piridava* 'I hand over', *podava* 'I obtain', *rozdava* 'I hand out', *vydava* 'I give away'. In Latvian Romani we find, based on *dža-* 'to go', *iedža-* 'to go in' and *piedža-* 'to approach', and in some Greek Romani dialects we find *dikh-* 'to see', and *ksanadikh-* 'to see again'.

Markers of adjective comparison in Romani often derive from pre-posed, unbound or semi-bound markers of the various contact languages, such as Slavic *po*, Romanian *mai*, and Turkish *daha* for the comparative, Slavic *naj*, Hungarian *leg-*, Turkish *en* for the superlative. A bound comparative/superlative marker *-eder*

had been borrowed into Proto-Romani from Iranian, and continues to be used in many dialects. Sinti Romani for example uses it for the comparative – thus *sik* 'fast', *sikedər* 'faster' – and adopts the German superlative, including both its synthetic ending and an accompanying preposition: *am sikestə* 'fastest', (dialectal) German *am schnell-ste(n)*.

There are no parallels to the borrowing of productive derivational morphology in Domari, possibly due to the highly inflectional character of Arabic, which makes it difficult to isolate individual morphemes as markers of a derivational procedure. The tendency is instead to borrow entire word-forms. This tendency has an interesting impact on the domain of adjective comparison. Arabic employs a morpho-phonological template *áCCaC* to derive comparative/superlative forms from consonantal roots: *kbīr* 'big', *ákbar* 'bigger'; *zġīr* 'small', *ázġar* 'smaller'. This template cannot easily be isolated or integrated into the agglutinative-inflectional morphological structure of Domari, nor is it simple or even possible to break down Domari adjectives such as *tilla* 'big' or *kištota* 'small' into tri-consonantal roots for insertion into the Arabic-based derivation template. The solution adopted by Domari speakers is to borrow the full Arabic word-form for all comparative/superlative forms, resulting in complete borrowing-based suppletion of the inventory of adjectives: *tilla* 'big', *ákbar* 'bigger'; *kištota* 'small', *ázġar* 'smaller'.

6. Other morpho-lexicon

This section surveys a number of categories that are typically expressed in both languages by unbound, often uninflected function words. Romani borrows its numerals *efta*, *oxto*, *enja* ('seven-nine') as well as higher numerals above twenty (with the exception of 'one hundred') from Greek. Many dialects tend to replace these higher numerals through loans from the contemporary contact languages. Domari speakers in Jerusalem are unable to recall non-Arabic numerals above 'five', with the

exception of 'ten' and 'hundred', although the full set of numerals is documented for speakers in the same community in the early 1900s (cf. Macalister 1914).

Some varieties of Sinti Romani have borrowed the dialectal German negator *nit*, which follows the finite verb, as it does in German: *me džinau nit* 'I don't know'. Other Sinti varieties use the German-derived emphatic particle *gar*, roughly 'indeed', as a negation particle: *me džinau gar* 'I don't know'. This functionalisation is internal to Sinti, and appears to have derived via the German expression *gar nicht* 'not at all'. Domari borrows the Arabic negator *miššš* in non-lexical predications – *pandži miš mišt-ēk* 'he *not* ill-COP.M.SG' = 'he is not ill'. The Arabic negator *mā ... -š* accompanies all Arabic-derived inflected verbs (modals and auxiliaries) in Domari: *pandži mā kānš mišt-ēya* 'he was not ill'.

Romani dialects tend to borrow prepositions that express more peripheral and more complex local relations. These include relations that have complex reference points, such as 'between', 'along', 'through', and 'around', those that involve separation from a source, such as 'from', 'toward', 'against' and 'since', and especially those that convey contrast with a presupposed set, such as privative 'without' and 'except', and replacive 'instead of'. Typical borrowings into various dialects include *pretiv/protiv* (Slavic) 'against', *is* (Slavic) and *fon* (German) 'from', *za* (Slavic) and *bis* (German) 'until', *de* (Romanian) 'since', *bez/brzo* (Slavic), *xoris* (Greek), *utan* (Swedish), and *oni* (German) 'without', *vmesto/namesto* (Slavic) 'instead', *osven*, *skluchenje*, *kromje* (Slavic), *in loc də* (Romanian) 'instead', and *ektos* (Greek) 'except for'. The preposition 'with' is also borrowed in some dialects, e.g. Sinti Romani *mit* from German, Greek Romani *me* from Greek. Domari, by contrast, borrows most of its prepositions from Arabic. Core prepositions borrowed from Arabic are *maʕ* 'with', *min* 'from', *baʕd* 'after', *ʕan* 'about', *ʕand* 'at', and *žamb* 'next to', in addition to *badāl* 'instead of', *bala* 'without', *bēn* 'between'. Some inherited (Indo-Aryan) adverbial expressions of local relations, such as 'above', 'outside', and 'inside', are nevertheless retained.

As discussed above, Romani dialects borrow modality word-forms such as Slavic and Romanian *musaj/musi-/muši-* and *triba/treba/trobu-* 'must', Slavic *mora* 'must' and *može* and *mog-* 'can', Turkish-derived *lazimi* and *medžburi* 'must', Swedish-derived *moste* 'must', German-derived *braux-* 'need', and Greek-derived *prep-* 'must' and *bor-* 'can' (see Matras 2002, Elšík & Matras 2006). Domari, too, borrows most of its modals and auxiliaries from Arabic – *lāzim* 'must' and *mumkin* 'it is possible', *kān* which indicates past-habitual, *ṣār* 'to begin', *baqi* 'to continue', *bidd-*'to want' – an exception being the verb *sak-* 'to be able to'.

All Romani dialects borrow indefinite markers, and many indefinite word-forms are also borrowed. Borrowed markers include Slavic *(v)sako*, *ni-*, *bilo-*, Romanian-derived *vare-*, Turkish-derived *hič*, *her*, *bazi*, Hungarian *vala-*, and more. Borrowed indefinite word-forms are numerous and include Slavic *ništo* 'something, nothing, anything', Polish *zawsze* 'always', Romanian-derived *mereu* 'always', Hungarian-derived *šoha* 'ever, never' and *mindig* 'always', Greek-derived *čipota* 'nothing' and *kathenas* 'somebody', and many more. Domari borrows the Arabic indefinite markers *kull* 'every', *ayy* 'any', and *wala* 'no, none', and combines them with the inherited Domari expression for person (*ekak*) and otherwise with Arabic expressions for thing, time, place, and so on. The word for 'always' is Arabic *dāʔiman*.

Romani interrogatives are generally retained from the Indic stock of forms, though occasional borrowings are found in particular among the interrogatives of quantity ('how much?'). Domari borrows both from Arabic – *qadēš?* 'how much /many?', and *waqtēš?* 'when?' – along with the determiner-interrogative *ayy?* 'which?'. Subordinating particles in Romani are generally grammaticalised interrogatives, but word-forms for 'because',

'although', and 'if' are frequently borrowed, as are, occasionally, relative pronouns and the factual complementiser, for which we often find Greek *oti*, Bulgarian *či*, Romanian-derived *ke*, Italian *ke*, and Hungarian-derived *hodž/hod/hoi*. Romani dialects always borrow 'but' from the contemporary or recent contact language (e.g. Slavic *no*, *po* and *ali/ale*, Hungarian *de*, Turkish *ama*, Greek *ala*, German *aber*). Many Romani dialects also borrow 'or' and 'and'. These latter two are borrowed frequently, but are sometimes retained from an older contact language. Ajia Varvara Romani in Athens for instance has *ja* 'oder' from its recent contact language Turkish, but *ala* 'but' from its current contact language Greek; Finnish Romani has *elle* 'or' from Swedish, but *mut* 'but' from Finnish; Manush Romani in France has German *un* 'and' and *otar* 'or', but French-derived *me* 'but', and so on (cf. again Matras 1998 & 2002, as well as Elšík & Matras 2006). Domari borrows all its connectivity devices from Arabic, including all coordinating and subordinating conjunctions, the relative particle *illi* and the factual complementiser *inn-*. Both languages always borrow discourse particles and interjections from the immediate contact language.

Finally, both languages borrow a considerable amount if not the full inventory of word-forms in the domains of phasal adverbs and focus particles. In Romani these are always loans from European languages, but they are relatively stable compared to other borrowed function words such as connectors or discourse markers. Thus some varieties of German Romani (Sinti) preserve Greek-derived *komi* 'still', but have German *schon* 'already' and *bloß* 'only', while Lovari Romani in Poland, Hungary, and Russia preserves Romanian-derived *inke* 'still', *aba* 'already', and *feri* 'only', and Burgenland Romani in Austria retains Hungarian-derived *meg* 'still', *imar* 'already', and *čak* 'only'. Domari borrows all its relevant particles – *bass* 'only', *kamān* 'too', *ḥatta* 'even', and so on – from Arabic.

7. Discussion

The present contribution was intended to address the question of which grammatical categories, especially in morpho-syntax and morpho-lexicon are more prone to contact-induced change, and which are more reliable indicators of genetic-historical inheritance - and so, in comparative perspective, of language-genetic relatedness. The comparative or historical linguist approaching Romani and Domari will find an entire set of categories thoroughly unreliable as indicators of inheritance, and at the same time diagnostic of recent and fairly recent contacts leading to the cross-linguistic diffusion of word-forms. This includes discourse markers, connectors, phasal adverbs, focus particles, indefinite expressions and indefinite markers, expressions of modality (modal verbs and modal particles), and the marking of comparison in adjectives. Expressions of aspect, nominal derivation, plural formation, prepositions, lower numerals, negation and interrogatives constitute a mixed set that is penetrated by borrowings to some extent. The reasons for the susceptibility of these categories to borrowing is as functional as it is, quite possibly, facilitated by their relative structural autonomy (see discussion in Matras 2007 & 2008).

By contrast, case and possessive inflection, demonstratives and personal pronouns, and verb inflection are based, by and large, on inherited material. The categories are not, however, impenetrable to borrowings. Nonetheless, borrowing in these categories is not just quantitatively restricted, but it is also limited to particular patterns. In case inflection, the only evidence for contact-induced change alluded to above was the semantic extension of an inherited inflectional marker due to its resemblance to a functionally related marker in the contact language (Zargari Romani *-e*, in contact with Azeri). It is crucial, at this point, to distinguish two types of contact-induced change. The first, to which most of the present discussion has been devoted, is the direct replication of linguistic Matter or

phonological shapes of morphs along with their meanings or aspects thereof. The second is the replication of Pattern – the linear arrangement of morphs, or the meaning associated with them (cf. Matras & Sakel 2007, Matras 2008). Since our focus is on identifying inherited linguistic material, patterns have been largely excluded from this discussion.

In terms of Matter-replication, we therefore find nothing in the domain of case inflection. In pronouns, we saw the borrowing of plural markers in several instances, and the borrowing of word-forms in the case of Molisean Romani 3PL and reciprocal 1PL, and the Domari reciprocal and reflexive. Interestingly, these are all parts of the pronominal paradigm that are generally volatile. Romani dialects frequently renew the form of the 3PL (cf. Elšík & Matras 2006), and reciprocals and reflexive are frequently composed of secondary expressions, and so constitute recent grammaticalisations. A 'core' consisting of the deictic (first and second person) pronouns remains, in this survey, immune to contact. Demonstrative stems are generally inherited (though they may be augmented by borrowed markers), but here too we find considerable internal renewal and a great diversity in the formation and re-formation of stems (cf. Matras 2002).

In verb inflection, finally, with the only exception of three borrowed person affixes in Slovene Romani (one of them derived from Greek), the distribution of borrowed markers is constrained by an internal compartmentalisation of the system. Thus Greek-derived tense markers may accompany borrowed word-stems in Romani, but they do not serve as independent markers of tense and do not compete with or replace the inherited tense markers. The Greek-derived 3SG marker *-i* is found only in (some) borrowed verbs (except in Slovene Romani, where it is generalised). Use of verb inflection from the contact languages is otherwise characterised by a wholesale import of the entire paradigm, rather than of individual markers, accompanying either just modal verbs (in Domari and some Romani dialects, such as Gurbet), or modal verbs and a selection of borrowed lexical verbs (as in Epirus

Romani), or all borrowed verbs from the particular contact language (as in some Romani dialects in contact with Turkish). In all cases, the imported verb paradigm exists in complementary distribution alongside the inherited inflection paradigm.

The one additional case of a change to individual markers (exemplified above, for Romani from Sindel) can be explained, much like the case of 3PL pronouns in some Romani dialects, as a replication of a plural morpheme rather than the substitution of an entire person marker. Adding both cases together – the replication of plural formation in pronouns and in the person concord markers on the verb – we might generalise that plural inflection is more prone to borrowing. Indeed, the employment of borrowed plural markers on nouns provides further confirmation.

Even within the categories that show relative resilience to borrowing, we therefore find some borrowing, albeit tightly governed by a number of constraints: plural inflection markers are more likely to be borrowed independently of other markers in the paradigm, and full person concord paradigms may accompany borrowed lexical verbs and especially modals but are not likely to diffuse into inherited verbs. Thus while no domain of grammar, save perhaps case inflection, appears to enjoy general immunity toward borrowing of linguistic matter, inflection does show considerable stability in the two languages examined, and deictic and anaphoric paradigms are volatile but their renewal is driven by internal pressures and draws on internal (inherited) resources. The historical linguist aiming to reconstruct the history of Romani or Domari – two languages that have been in intense contact, as low-prestige, minority languages spoken exclusively by bilinguals – would encounter on the whole coherence between the source of core vocabulary and that of inflectional morphology, with disruptions being by and large recognisable as wholesale imports of paradigms along with borrowed lexicon.

Nonetheless, we must bear in mind that this finding draws on two important factors. The first is the relatively shallow time depth of contact developments. The radical compartmentalisation

in the verb system of some Romani, not to mention the complete replacement of the inventory of expressions in the more 'borrowable' categories, represents a time-depth of merely 500-600 years. The Greek impact on Romani is much older in terms of its absolute age, but considering that it receded with the breakup of the Byzantine Empire, we might again reconstruct a period of contact of somewhere between 400-500 years. Similarly, Domari has been in contact with Arabic possibly since the Saladin conquests in the twelfth century at the very earliest and in all likelihood for a much shorter period. Now, there is no reason to believe that contact-induced change should proceed at a steady pace, and so we should not assume that, had contact between Romani and Greek continued, then substantial additional components of the language might have changed. Indeed, we may gain some insights by examining those Romani dialects that are still spoken in Greece today. Nonetheless, both our sample languages represent periods of contact that are not immensely long; we cannot rule out the possibility that developments extending for millennia might have a stronger impact on a language.

The second factor is the nature of contact and its relevance to language maintenance. Both Romani and Domari are 'languages in contact', rather than 'contact languages'. This means that the contact influences that they show are a result of prolonged and gradual change. Even a comparison of Domari sources from the early twentieth century (Macalister 1914) with material collected recently in the same community (Matras 1999) shows some degree of gradual change among generations (in this case, the loss of inherited higher numerals during this time interval). The outcome can be very different in cases of an abrupt turnover of the balance of functions between the languages of a bilingual community, and the partial abandonment of a language within one or two generations. This is the case with Angloromani, the variety of British Romani that ceased to exist as an everyday family language toward the end of the nineteenth century, and was replaced instead

by a register that serves the function of an emotive discourse mode (cf. Matras et al. 2008). This functional turnover carries with it the abrupt loss of most grammatical inflection paradigms, especially finite verb inflection, and the reduction of the old language to, by and large, just a lexical inventory. Similar processes are attested from contemporary, ongoing cases of language shift, sometimes involving the partial retention of grammatical material (cf. O'Shannessy 2005, McConwell & Meakins 2005). It is now well-established in the contact linguistic literature that such cases of Mixed Languages (Bakker 1997, 2000) are not reliable indicators of genetic inheritance (cf. Thomason 1997). In the present contribution, I hope to have been able to illustrate that in languages that are not contact languages of the latter type, but have nevertheless been shaped by the huge impact of contact, a core of coherent linguistic-structural material representing an unambiguous source of genetic inheritance can still be identified.

Endnotes

[1] In Domari, this set also attaches to nouns to express possessive relations: *kury-om* 'my house', *kury-or* 'your house', *kury-os* 'his/her house', etc.
[2] See Matras (2002) for an introduction and overview; for general information, bibliography, and on online database comparing Romani dialects in transcription and sound see the Manchester Romani Project website on http://romani.humanities.manchester.ac.uk/

References

Baghbidi, H. R.
2003 The Zargari language. An endangered European Romani in Iran. *Romani Studies* 5/13: 123-148.

Bakker, P.
1996 Language intertwining and convergence: typological aspects of genesis of mixed languages. *Sprachtypologie und Universalienforschung* 49: 9-20.
2000 Rapid language change: Creolization, intertwining, convergence. In: Renfrew, C., McMahon, A. & Trask, R. L. (eds), 585-620.

Elšík, V. & Matras, Y.
2006 *Markedness and language change: The Romani sample.* Berlin: Mouton de Gruyter.

Elšík, V.
2007 Grammatical borrowing in Hungarian Rumungro. In: Matras, Y. & Sakel, J. (eds), 261-282.

Field, F. W.
2002 *Linguistic borrowing in bilingual contexts.* Amsterdam: John Benjamins.

Grierson, G. A.
1922 *Linguistic survey of India.* Calcutta: Superintendent Government Printers.

Hancock, I. F.
1995 On the migration and affiliation of the Ḍōmba: Iranian words in Rom, Lom, and Dom Gypsy. In: Matras, Yaron. (ed.), 25-51.

Macalister, R. A. S.
1914 The language of the Nawar of Zutt, the nomad smiths of Palestine. *Gypsy Lore Society Monographs 3.* London: Edinburgh University Press.

Masica, C. P.
1991 *The Indo-Aryan languages.* Cambridge: Cambridge University Press.

Matras, Y. & Sakel, J.
2007 Investigating the mechanisms of pattern-replication In Language convergence. *Studies in Language 31*: 829-865.

Matras, Y. & Sakel, J. (eds)
2007 *Grammatical borrowing in cross-linguistic perspective.* Berlin: Mouton de Gruyter.

Matras, Y.
1998 Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36: 281-331.
1999 The state of present-day Domari in Jerusalem. *Mediterranean Language Review* 11: 1-58.
2002 *Romani: A linguistic introduction.* Cambridge: Cambridge University Press.
2007 The borrowability of grammatical categories. In: Matras, Y. & Sakel, J. (eds), 31-74.
2008 *Language contact.* Cambridge: Cambridge University Press.

Matras, Y. ed.
1995 *Romani in contact: the history and sociology of a language.* Amsterdam: Benjamins.

Matras, Y., Gardner, H., Jones, C. & Schulman, V.
2008 Angloromani: A different kind of language? Anthropological Linguistics.

McConvell, P. & Meakins, F.
2005 Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25: 9-30.

O'Shannessy, C.
2005 Light Walpiri: A new language. *Australian Journal of Linguistics* 25: 31-57.

Renfrew, C., McMahon, A. & Trask, R. L. (eds)
2000 *Time depth in historical linguistics.* Vol. 2. Cambridge: McDonald Institute for Archeological Research.

Sampson, J.
1923 On the origin and early migrations of the Gypsies. *Journal of the Gypsy Lore Society*, third series, 2: 156–169.

Thomason, S. G.
1997 A typology of contact languages. In: Spears, A. A. & Winford, D. (eds) *The structure and status of pidgins and creoles.* Amsterdam: John Benjamins. 71-88.

Turner, R. L.
1926 The position of Romani in Indo-Aryan. *Journal of the Gypsy Lore Society*, third series, 5: 145-189.

Windfuhr, G.
1970 European Gypsy in Iran: A first report. *Anthropological Linguistics* 12: 271-292.

Iranian Archaisms vs. Vedic Innovations – and the Indo-Iranian Unity

Rüdiger Schmitt
Laboe (Germany)

The author's main purpose is to survey as fully as possible the archaic traits of Iranian which are in sharp contrast with secondary changes, i. e. obvious innovations in Old Indo-Aryan (even in Vedic times) and by this are proving that all Indo-Aryan linguistic evidence first must be compared with the Iranian data alone for reconstructing Proto-Indo-Iranian, since only that unattested comon Indo-Iranian proto-language (the most important features of wich are listed here as well) is on the same level as the other members of the Indo-European language family. Just the archaism of Old Avestan (the language of Zarathushtra's songs) makes it plainly clear that the Old Indo-Aryan language is not the only solid fundament found within the Indo-Iranian area and is not so close to the Indo-European proto-language that one may pass over all differences attested elsewhere in silence.

The[1] family of the Iranian languages and the Indo-Aryan language family are so closely related to one another, as is well known, that they are rightly traced back to a common source[2], i. e. a common stage of the linguistic development in the period after the disintegration of the Indo-European proto-language, which came to an end, when the (Proto-)Indo-Aryans moved away towards the Indian subcontinent. Because of this great affinity of the Old Iranian and the Old Indo-Aryan languages in their earliest records one often tried to illustrate this by transposing Avestan sentences in accordance with the phonological correspondences into Vedic. Just as well we may reconstruct by means of the comparative method this common source itself, the Proto-Indo-Iranian language, even if this language certainly was not absolutely homogeneous, but in all probability was rather a continuum of mutually understandable dialects[3].

This intermediate proto-language or (in German) *Zwischengrundsprache*, as one could say[4], is defined by a large number of phonological, morphological, syntactical, lexical, and phraseological isoglosses found exclusively in both these language-families, including the speakers' self-designation as „Aryan" (Ved. *ā́r(i)ya-*, YAv. *a*^{i}*riia-*, OP. *a-r-i-y-* /ariya-/, etc.). From the ancient texts it is clear that this term „Aryan" originally is an ethnonym characterizing the next larger group after the Persians (in the inscriptions of the Achaemenid kings) or the whole of the national community (when in Yašt 5, 69 Jāmāspa speaks of „all the other Aryans", *vīspe aniie a*^{i}*riie*) and the Aryan population of India (*ā́ryaṃ várṇaṃ*, RV 3, 34, 9d) in contrast to the non-Aryan (*dā́saṃ várṇam*, RV 2, 12, 4b) respectively. But at least in an Old Persian and a Bactrian inscription we find this term also as the name of the language in question. The importance of this name had been stressed already by Meillet (1908: 24), for whom it was the only direct testimony of any such linguistic unity from the period after the disintegration of Proto-Indo-European.

Notwithstanding this clear testimony, the term „Aryan" has often been used, in particular during the 19th century, to denote the language of the undivided Indo-European linguistic community in general, whereas for others, and especially for Indian scholars, it is even today more or less synonymous with the Vedic people and its language. For this reason and in order to avoid any misunderstanding I prefer here the term „(Proto-)Indo-Iranian" to „(Proto-)Aryan" as the name of the common source of both the Indo-Aryan and the Iranian languages.

The phenomena relevant for postulating such linguistic subgroups, which Hock (2000: 122 f.) compared with what in textual criticism are hyparchetypes, are according to Leskien's principle only exclusive common linguistic innovations of the languages involved. And these innovations should not concern trivial points such as the levelling of irregularities by analogy and should not be easily borrowed. Moreover, one must be able to rule out independent parallel development, e. g. by means of so-called

„cumulative evidence". Among the most important linguistic facts constituting this (Proto-)Indo-Iranian linguistic unity are the following: (1) the effect of the law of the palatals (PIE. *ke > IIr. *ce > *ca vs. PIE. *ko > *ka, etc.); (2) the merger of PIE. *a, *e, *o into IIr. *a (and of *ā, *ē, *ō into *ā); (3) the lengthening of PIE. *ŏ to IIr. *ā in open syllables according to Brugmann's law (cfr. PIE. *dóru [Gk. δόρυ] > IIr. *dā́ru > Ved. dā́ru, YAv. dā^{u}ru); (4) the sonorization of the so-called laryngeals PIE. *h_1, *h_2, *h_3 in certain phonotactic positions and varying range (cfr. PIE. *p$h_2$tér- „father" > Ved. pitár-, Av. pitar-, but dat. sg. *p$h_2$tr-éi̯ > Ved. pitré, OAv. f^{ə}δrōi, etc.); (5) the effect of Bartholomae's law (IIr. *b^{h}t > *b^{h}d^h > *bd^h, etc.), which for other scholars is of Proto-Indo-European date, however; (6) the spread of the case ending IIr. *-nām in the gen. pl. of vocalic stems, which was not generalized, however, in Avestan i- and u-stems[5]; (7) the creation of a passive voice conjugation by a present stem in IIr. *-i̯a- and of a future stem in IIr. *-si̯a-; (8) the change of the initial *ḱ- of PIE. *ḱḗrd-/*ḱr̥d- „heart" (cfr. Gk. κῆρ, Lat. cord-, etc.) into IIr. *ǰ^h- (*ǰ^{h}ŕ̥d- > Ved. hŕ̥d-, OAv. zər^{ə}d-) probably in analogy to some synsemantic word; etc. etc.[6]. Other characteristics of Proto-Indo-Iranian include (9) the survival of the Proto-Indo-European laryngeals, which is proven e. g. by cases of hiatus in Vedic and Avestan metrics or by the different development of PIE. *r̥H, *l̥H to Ved. ir/ur before vowel and īr/ūr before consonant, but equally ar in (Proto-)Iranian (cfr. PIE. *pl̥h_1-ú- „much" > IIr. *pr̥Hú- > Ved. purú-, fem. pūrvī́-, but OP. paru-, etc.)[7]; (10) the reorganization of the ā-declension after the ī-stems, and (11) the origin of complex paradigms of demonstrative pronouns, in which various stems are combined, as we see it in Ved. ay-ám, fem. iy-ám, ntr. i-d-ám, acc. masc. imá-m, etc.

All in all, phonological criteria are the most important anyway, particularly in case that the change observed leads to an irreversible result. For Proto-Indo-Iranian there is also the fact that we find a great number of lexical correspondences (often

concerning religion and mythology) and just as well even more complex idiomatic usage attested in both the oldest Vedic and Avestan texts. It may be sufficient, to quote here Ved. uttānáhasta- námasā (RV 6, 16, 46b, etc.) „with hands outstretched in reverence" = OAv. ustānazasta- nəmaŋhā (Y. 28, 1a) and Ved. hr̥dā́ mánasā (RV 1, 61, 2c, etc.) „with heart (and) thought" = OAv. zərədā(cā) manaŋhā(cā) (Y. 31, 12b). And if a singer calls his song a product „without a former (model)" (Ved. ápūrvyam = OAv. apao^{u}ruuīm), we can infer from this formulation that there have been previous singers and songs, poets and poems, i. e. a poetic tradition at least from common Indo-Iranian times and thus a hereditary poetic language handed over from generation to generation. All this is clear evidence not only for a close relationship of the languages, but also for an intellectual and cultural community of the speakers of that language.

In consequence of those old traditions it is small wonder that one finds often highly archaic linguistic features in the entire Indo-Iranian world. The complex interrelation of mobile accent and ablaut in the root, the stem-formation and the case-ending recognized in the second half of the 20th century as the fundamental primary system operating in noun morphology (instead of the traditional classes [and subclasses] of a-, ā-, o-, i-, ī-, u-stems, etc.) is most plainly preserved in Indo-Iranian, e. g. in the two classes of ī-stems with or without ablaut (the stems of the devī- and the vr̥kī-type): cfr. (with ablaut) Ved. nom. sg. nā́r-ī „wife", gen. nā́r-yā-ḥ = Av. nā^{i}rī́, gen. nā^{i}riiå (< *-i̯āh) in contrast to (without ablaut) Ved. vr̥k-ī́-ḥ „she-wolf" (< PIE. *-íh₂-s), gen. vr̥k-i̯y-àḥ (< *-íy-as < PIE. *-íh_2-os) = Av. °īš, gen. °iiō. Parts of the highly archaic inflexion of PIE. *$h_2$éu̯s-ōs „dawn" have also survived into Proto-Indo-Iranian and its oldest continuants, even if the root-ablaut has been levelled: Ved. nom. sg. uṣā́ḥ = OAv. ušå < IIr. *Huš-ā́s for PIE. *$h_2$éu̯s-ōs-∅ (cfr. Gk. ἠώς, Aeol. αὔως), acc. Ved. uṣā́s-am = YAv. ušåŋhəm for older Ved. uṣás-am < IIr.

*Huš-ás-am for PIE. *$h_2$éu̯s-os-m̥ (cfr. Gk. ἠῶ < *ἠόα), gen. Ved. uṣáḥ < *uṣ-ṣ-ás < PIE. *$h_2$us-s-és.

Regardless of this archaism of Indo-Iranian in general it is a matter of fact that Proto-Indo-Aryan and Proto-Iranian are different in one respect or another, i. e. by varying innovations in only one branch. At this occasion I should remark that I see no reason to question or even to deny, as Tremblay 2005b did it, the existence of Proto-Iranian as the dialectally more or less uniform common ancestor of all Iranian languages. On the one hand his argument that most of the typical Iranian innovations are missing in one or another dialect, as an *argumentum e silentio* does not mean very much, and on the other hand there is, apart from others, one Iranian innovation so much distinctive and remarkable in form and function that it is proof enough for claiming a Proto-Iranian language. I speak of the development of an own mood „frequentative“ indicating repeated habitual actions of the past, which by its form is a present optative often with augment *a*- and thus looks like an imperfect optative[8]. Obviously it was at a quite early date that Proto-Iranian had split off from Proto-Indo-Iranian, so that the development of more or less far reaching innovations is not surprising at all.

But the aim of this paper is to concentrate on the archaic traits found in the Iranian languages and especially in (Old) Avestan, since all in all Avestan is so an archaic language as Rigvedic Indo-Aryan and Homeric Greek. That the Iranian branch of the Indo-Iranian language group[9] often reveals more archaic traits than the Indo-Aryan[10] can be seen quite clearly in phonology[11]. Most characteristic for Indo-Aryan (as for all the languages spoken on the Indian subcontinent, no matter of what language family) are the cerebral or retroflex phonemes (*ṭ*, *ṭh*, *ḍ*, *ḍh*, *ṇ*, *ṣ*) that are unknown to the entire rest of Indo-European except some modern Iranian languages (like Pashtō) which have been influenced by neighbouring Indo-Aryan and/or Dravidian languages; therefore those sounds with certainty are an exclusive innovation on the part of Indo-Aryan already in pre-Vedic times[12].

One particular case is the development of a retroflex (combined with compensatory lengthening of a preceding short vowel) in clusters like IIr. *žd$^{(h)}$ as, e. g., in Ved. *mīḍhá*- „reward, prize“ from IIr. *mižd^{h}á- (< PIE. *mizd^{h}ó-, cfr. Gk. μισθός, etc.) as opposed to PIr. *mižda- (> OAv., YAv. *mīžda*-), since Indo-Aryan in general is characterized by the loss of the voiced sibilants IIr. *z and *ž (cfr. also PIE. *m̥nz-d^{h}éh_1- > IIr. *mazd^{h}ā́- > Ved. *medhā́*- „wisdom“, but OAv. etc. *mazdā*-), owing to which loss the Indo-Aryan forms as such have become opaque, i. e. totally incomprehensible[13]. A good example for illustrating the archaism of Iranian can be seen also in the various counterparts of the cluster OIA. *kṣ*, in which a number of distinct (combinations of) consonants have fallen together that are still clearly distinguished from one another in Iranian: cfr. Ved. *dákṣiṇa*- „right(-hand)“ vs. YAv. *dašina*- < IIr. *dáćšina- < PIE. *déḱsino-; Ved. 3 sg. aor. subj. act. *vakṣat* (from *vah* „to drive, draw“) vs. YAv. °*uuažat̰* < IIr. *u̯aȷ́^{h}sat < PIE. *u̯eǵ^{h}set; Ved. *kṣatrá*- „sovereignty“ vs. OAv., YAv. *xšaϑra*- < IIr. *kšatrá-; Ved. *kṣám*- „earth“ vs. OAv., YAv. *zam*- < IIr. *dȷ́^{h}ám- < PIE. *d^{h}ǵ^{h}ém-; or the root Ved. *kṣar* „to flow“ vs. YAv. *γžar* < IIr. *dȷ́^{h}ar < PIE. *d^{h}g^{uh}er. All these examples that show the effect of irreversible sound changes make it clear that the Indo-Aryan forms cannot be the starting-point for any reconstruction of the situation in the proto-language, since on the basis of them alone one could not predict the Iranian equivalents or those in other languages (as, e. g., Greek) where likewise different clusters occur.

A list of those Iranian archaisms that are faced with secondary changes and therefore obvious innovations on the Vedic side[14] to the best of my knowledge has never been drawn up, although this would have been (and is) of essential importance for proving (1) that Indo-Aryan (with its variants Vedic and Sanskrit) is not *per se* the most ancient and most archaic of the Indo-European languages and the one most important for reconstructing the proto-language and (2) that the Indo-Aryan language and culture must have immigrated into India and do not

originate there. So I will catch up on this task and try to list the most important Vedic innovations or Iranian archaisms in a more systematic manner[15], even if restricted to Old Iranian[16] (esp. [Old] Avestan as its most archaic representative).

The Old Avestan language, as used by Zarathushtra in his *Gāϑā*, without doubt was to a certain extent an artificial language, in which various phenomena had run together, not necessarily phenomena of different dialects, since just as well archaic forms may be found beside younger ones, those of normal language beside those of poetic diction, if not even later developments arosen only in the long course of the oral transmission. Of some importance are also the changes and standardizations made by diascevasts (e. g. the repetition of preverbs directly before the verb in cases where they originally were separated from it by tmesis), which make it perfectly clear that the text of our manuscripts, even if it is rather correct, is not by any means the original text itself. The supposed traces of younger innovations in the language of the *Gāϑā* had been discussed in an influential study by Meillet (1917: 183–195) which, however, was closely examined and rightly criticized recently by Kellens 2006. In rather faithful continuation of the Indo-Iranian poetic tradition the Old Avestan language unlike Younger Avestan and Old Persian[17] is as highly archaic as the language of the oldest collection of Vedic hymns, the *Rigvedasaṃhitā*, and indeed it often is more archaic than the Rigvedic language. Obviously this has to do with the fact that, as we can see from the vocabulary, the influence of the indigenous pre-Aryan population was much smaller in Iran than in India.

Phonology:

(1) Laryngeal hiatus, more exactly: hiatus caused by the disappearance of an intervocalic laryngeal, can be observed (at least by the effects caused by it) in Old Avestan much more faithfully than in Rigvedic and actually it is strictly observed, as we can see from long vowels resulting from PIE. **-VHV-* which in part are measured as two syllables according to the metre and

therefore cannot be contracted as usual: cfr. OAv. nom. *mazdā̊* „Mazdā“ (repeatedly disyllabic, < PIr. **mazdāh* < IIr. **mazdaH-s*) vs. gen. *mazdā̊* (repeatedly trisyllabic, < IIr. **mazdaH-as*); OAv. 2 sg. aor. inj. act. *dā̊* „you gave“ (repeatedly monosyllabic, < IIr. **daH-s*) vs. nom. sg. *dā̊* „gift“ (Y. 28,6a disyllabic, < IIr. **daH-as* < PIE. **déh*$_3$*-os*); OAv. *vāta-* „wind“ (Y. 44, 4d trisyllabic) = Ved. *vā́ta-* (repeatedly trisyllabic) < IIr. **u̯áHata-* < PIE. **h*$_1$*u̯éh*$_2$*-n̥to-*. Not only those relics – the entire Gathic evidence is discussed by Monna (1978: 97–103) – show the great importance of the highly archaic Old Iranian languages for the laryngeal theory (see Mayrhofer 2005: 117 ff., quoting once more the words of Benveniste[–Renou] 1934: 182: „le témoignage védique vaut par sa richesse, le témoignage avestique par sa fidélité“).

(2) The varying development of the laryngeal PIE. **h*$_2$, which in the hysterokinetic word for „father“ has been partly vocalized to *i* and partly disappeared in Iranian, has produced an extremely „irregular“ paradigm. Whereas in Vedic the old inherited ablaut has been levelled and *-i-* instead of *-Ø-* has been generalized by analogy, forms with or without *-i-* have been preserved side by side in Avestan alone:
sg. nom. OAv., YAv. *ptā*, OAv. *tā*[18], YAv. *pita* vs. Ved. *pitā́* < PIE. **ph*$_2$*tḗr*,
sg. acc. OAv. *p*^{a}*tarə̄m* /ptar°/[19], YAv. *pitarəm* vs. Ved. *pitár-am* < PIE. **ph*$_2$*tér-m̥*,
sg. dat. OAv. *f*^{ə}*δrōi* /fθrai̯/, *piϑrē*, YAv. *piϑre* (by analogy) vs. Ved. *pitr-é* < PIE. **ph*$_2$*tr-éi̯*,
pl. nom. YAv. *p*^{a}*tarō*[20] vs. Ved. *pitár-as* < PIE. **ph*$_2$*tér-es*,
pl. dat. YAv. *ptər*^{ə}*biiō* vs. Ved. *pitr̥̄-bhyas* < PIE. **ph*$_2$*tŕ̥-bh°*.
Even if it is absolutely clear that the irregular Avestan paradigm is more genuine than the Vedic with its constantly recurring *pit°*, we can only voice suppositions about the exact forms of the Indo-Iranian stage. We cannot be sure where to put IIr. **pHt°* with consonantal *H* and where the form with its vocalized allophone, since the exact conditions (presumably those of accentual nature) are unknown, under which the laryngeals had been vocalized; as

Tichy (1985: 231–236) has shown, in all probability the vocative case PIr., IIr. **pítar-* (= Ved. *pítar*, Gk. πάτερ < PIE. **pḥ*$_2$*ter*) is the only source for this (in contrast to IIr. **pHtā́*, **pHtáram* etc.) and all other **pit°* forms in Avestan and Old Persian (as in Vedic) must be of secondary origin.

(3) All Avestan case-forms, compounds and derivatives with the weak stem of the numeral *caϑβārō* „four" present without exception *catur°* before vowel (acc. *caturə̄*, gen. *caturąm*) and *caϑru°* before consonant (*caϑru-dasa-* „14th", *caϑru-š* „four times", *caϑrušum* < **caϑru-šu̯a-m* „one fourth", *caϑru-gaoša-* „four-eared", etc.), whereas Vedic has only *catur-* in acc. *catúras*, instr. *catúrbhis*, *cáturdaśa* „14", *catús* „four times", *cátuṣ-pā̆d-* „four-legged", etc., obviously owing to a generalization of the antevocalic variant. That the Avestan doublet *catur°/caϑru°* represents the original state of IIr. **čatur*-V vs. **čatru*-C, is confirmed beyond any doubt by other reflexes of PIE. **k̯ᵘetru*-C (< **k̯ᵘetu̯r-* by metathesis already in the proto-language) in Lat. *quater*, *quadru-*, Gaul. *petru-*, etc.

(4) Old Indo-Aryan presents a number of words, in which an original **s* has been assimilated to another sibilant: cfr. well-known examples like Ved. *śaśá-* „hare" < **śasá-* = Av. *saŋha-* (in Pahlavi writing *s'ng*) = Khot. *saha-* < IIr. **ćasá-* < PIE. **ḱasó-* or Ved. *śváśura-* „father-in-law" < **sváśura-* = YAv. *xᵛasura-* < IIr. **su̯áćura-* < PIE. **su̯éḱuro-* (cfr. Gk. ἑκυρός, etc.). Since owing to the change of IIr. **s* > PIr. **h* the conditions were not the same in Iranian, such an assimilation could not happen there, and the forms look more archaic even later.

Morphophonology:

(5) Vr̥ddhi derivatives from basic words with *i* or *u* in the initial syllable in Avestan show without exception short diphthongs („guna") **ai̯*, **au̯* (written *aē*, *ao/əu*) contrary to Vedic long diphthongs („vr̥ddhi") *ai*, *au* < IIr. **āi̯*, **āu̯*: cfr. *haomanaŋha-* „benevolence" from *hu-manah-* vs. Ved. *saumanasá-*, *haosrauuaŋha-* „good name" from **hu-srau̯ah-* vs. Ved. *sauśravasá-*, *vaēδaiianā-* „watchpost" from IIr. **u̯i-dʰai̯Hanā-* „looking to and fro". Obviously the Avestan short diphthongs are older, particularly since corresponding relics like Ved. *bheṣajá-* „healing; medicine" = YAv. *baēšaza-* from *bhiṣáj-* „physician" are found (see Darms 1978: 367–375). The only such case in Old Persian is the month name *Θāigraci-*, a vr̥ddhi derivative from **Θigra-ka-* (perhaps the name of a garlic [harvest] festival; see Schmitt 2003: 36–39), which looks like exhibiting a long diphthong of the Indo-Aryan type; one may consider the spelling *ϑ-a-i°* instead of expected **ϑ-i°* /Θai̯°/, however, also as purely graphic and as a case of disambiguating (like gen. *Cišpāiš*, where **Cišpaiš* is expected).

(6) The exact retention of the inherited inflexional ablaut in root nouns is also in contrast to Vedic, where the paradigmatic ablaut has been abandoned and the lengthened grade has been generalized (as in Latin): cfr. Ved. nom. sg. *vā́k* „speech, voice", acc. *vā́c-am*, gen. *vāc-ás*, instr. *vāc-ā́*, nom. pl. *vā́c-as*, acc. *vā́c-as*, gen. *vāc-ā́m*; in Avestan the difference between the strong cases nom. sg. OAv., YAv. *vāxš*, acc. OAv. *vāc-əm*, YAv. *vāc-im*, nom. pl. YAv. *vāc-ō* and the weak cases (all YAv.) gen. sg. *vac-ō*, instr. *vac-a*, acc. pl. *vac-ō*, gen. *vac-ąm* is plainly recognizable. An archaism of the external sandhi is the sigmatic form *vāx-š* (= Lat. *uōx*)[21], the phonetic reality of which is supported by the Manichaean term (MP., Parth., Sogd.) *vāxš*. Another case of inherited ablaut within a root noun paradigm is in Av. *ziiam-* „winter" with acc. sg. YAv. *ziiąm* (after which nom. *ziiā̊*; cfr. *mazdā̊* vs. *mazdąm*, etc.) vs. gen. OAv. *zim-ō*, YAv. *zəm-ō*, which go back to IIr. **ǰʰi̯ám-* vs. **ǰʰim-*, the counterpart of Lat. *hiem-s* (via PIE. **ǵʰi̯ém-*). The Vedic language has no traces of this stem, but only thematic *himá-* „cold, frost", which probably is a back-formation based on compounds like YAv. *hazaŋrō.zəma-* „period of 1000 winters" or Ved. *(śatá-)hima-*.

(7) The Indo-Iranian word for „path" (as opposed to „road") is a formation with the laryngeal suffix PIE. *-*oh*$_2$- and characterized by ablaut alternation of both root and stem vowel: PIE. nom. sg. **pént-oh*$_2$*-s*, but gen. sg. **pn̥t-h*$_2$*-és* with double

zero-grade, when in this amphikinetic paradigm the accent fell on the case-ending. In Indo-Iranian times those zero-grade forms were affected by the aspiration of PIE. *th_2 to IIr. *t^h, which lead to, i. a., nom. sg. *pántās and gen. *pat^{h}ás respectively. One can easily imagine that such a paradigm with the two variations of IIr. *an vs. *a and *t vs. *t^h demanded to be evened out. Just this happened in Vedic, where we find nom. pánthās, gen. pathás etc., whereas Avestan preserved the original state unchanged: nom. YAv. *paṇtā̊*, acc. *paṇtąm* (cfr. Ved. *pánthām*), but abl. OAv. *paϑō* (= Ved. *pathás*), gen. YAv. *paϑō*, loc. OAv. *pa^{i}ϑī* (= Ved. *pathí*) etc., even instr. pl. OAv. *pad^{ə}bīš* vs. Ved. *pathíbhis* (in the end < PIE. *pn̥t-h_2-b^{h}ís). Since any parallel to this highly irregular type of inflexion is missing, it can be explained only as an archaism; and we learn from this also why the Indo-Iranian forms and among them the Vedic in particular differ from Gk. πόντ-ος, Lat. *pont-* and the others. Moreover the fact is of some importance that this variation of OIr. *pant-/*paϑ- is reflected by isolated Middle and New Iranian material, too (cfr. Mayrhofer 1996: 81 f., with lit.), indirectly also by Scythian, where *pant- is present in the hydronym Παντικάπης and *paϑ- maybe in the original of the loan-word E. *path*, NHG. *Pfad*, etc. (cfr. Mayrhofer 2006: 24).

(8) Because of the ablaut variation in the stem vowel of the (PIE.) *-es-/-os* neuters that show PIE. *-os* in nom.-acc. sg. and *-es- in all other cases, i. e. the type well-known from Gk. μένος, (Hom.) μένεος, Lat. *genus* (< *-os), *gener-is* (< *-es-es) and OCS. *nebo*, *nebese* (< *-es-es), we have to expect for stems ending in a velar stop that in (Proto-)Indo-Iranian (where PIE. *a, *e, *o actually merged into *a) this velar stop is preserved unchanged in nom.-acc. sg. and underwent palatalization according to the law of the palatals before original PIE. *-es-; thus, e. g., PIE. *$h_2$éu̯gos „power, strength" (cfr. Lat. *augus-tus* „majestic, solemn") > IIr. *áu̯gas, but PIE. *$h_2$éu̯ges- > IIr. *áu̯ǰas-. This opposition reflecting the old Proto-Indo-European ablaut pattern is found indeed in OAv. nom. sg. *aogō* (< PIr. *au̯gah) vs. instr. sg. *aojaŋhā* (< PIr. *au̯ǰah-ā), whereas in Vedic the palatalized stem

variant *ójas-* has been generalized throughout (cfr. nom.-acc. sg. *ójas*, instr. *ójas-ā*, etc.) as it is the case also in Younger Avestan, where nom.-acc. sg. *aojō* is attested, too. The exact correspondence between derived Ved. *ójas-vant-* „powerful" and OAv., YAv. *aojaŋ^{v}haṇt-* (< PIr. *au̯ǰah-u̯ant-) with *-ǰ-* in both forms is all right, since as it becomes quite clear from Greek (cfr., e. g., τελήεις „complete" < *teles-u̯ent-s to τέλος) we have to start from PIE. *°es-u̯ent-. Usually, however, as in the case of *ójas-*, only one of the two variants, either that with the velar or that with the palatal, is found through the entire paradigm. Other examples in which such analogical levelling of the velar/palatal opposition did not happen in (Old) Avestan, exist in a number of forms: e. g. in the verbal root *gam* „to go, come", where several *g°* forms survived, even if *j°* spread nearly everywhere, so that the situation is almost the opposite of that attested in Vedic, where only *g°* is found with at best one quite marginal exception in the compound *jamád-agni-* (if originally this is „going to Agni"); in the root *kar* „to do, make", where isolated *c°* forms are attested in the athematic root aorist (OAv. 3 sg. inj. *cōr^{ə}t̰* < *čar-t [cfr. Ved. *kar*], 3 sg. subj. *car-at̰* [cfr. Ved. *kárat*])[22], whereas Old Indo-Aryan has only *k°* forms; in the pronominal stems *ka-/ca-/ci-* (see under (18) below); or, scarcely to be recognized, in dat. sg. *haše* < PIr. *hač-i̯-ai̯ vs. nom. sg. OAv. *°haxā*, YAv. *haxa* „friend, companion", whereas the Vedic paradigm with nom. *sákhā*, dat. *sákhye* (< PIE. *sék^u-h_2-ō(i̯), *sék^u-h_2-i̯-oi̯) etc. has abolished that difference.

Morphology:

(9) The case ending of the dat. sg. of *a*-stems Ved. *-āya* = OAv. *-āiiā*, *-āi.ā* (cfr. OAv. *ahurāi.ā* „to the lord" = Ved. *ásurāya*) obviously is an Indo-Iranian innovation based on the inherited ending IIr. *-ai̯ < PIE. *-ōi̯ (cfr. Gk. -ῳ, OLat. *-oi*, etc.). This old *-āi̯ is attested in Avestan, too (cfr. OAv., YAv. *ahurāi*, etc.), as an archaism, whereas in Vedic it has been totally displaced by the

longer form. It must be remarked, however, that IIr. *-*āi̯a* itself was formed only in the Indo-Iranian period.

(10) In hysterokinetic *u*-stems only the Old Iranian languages (both Avestan and Old Persian) have preserved original forms of the strong stem to a larger extent: cfr. nom. sg. Av. °*bāzāuš* (e. g., OAv. *darəgō.bāzāuš* „long-armed"), OP. *dahyāuš* „land, province"; acc. sg. YAv. *nasāum* „corpse", OP. *dahyāum* (and *dahyāv-am*); nom. pl. YAv. *daŋ́hāuuō* = OP. *dahyāva*, YAv. *nasāuuō*; nom.-acc. du. YAv. *bāzauua* „both arms". The corresponding dual form *bāhávā* (= YAv. *bāzauua*) is the only relic at all, however, of the old inflexion in Vedic (cfr. Kuiper 1942: 40–45). A similar case is found among the hysterokinetic *i*-stems: Whereas both branches of Indo-Iranian agree in reflecting that inflexional type in Ved. *sákhā* „friend, companion" and OAv. °*haxā*, YAv. *haxa* (consider acc. sg. OAv. °*haxāim* = Ved. *sákhāyam*, dat. *haše* [< PIr. **hač-i̯-ai̯* with palatalization; cfr. (8) above] ~ *sákhye*, nom. du. *haxaiia* = *sákhāyā*, nom. pl. *haxaiiō* = *sákhāyas*), Av. *kauui-* (more exactly *kauuāii-*) „seer, prince" in sharp contrast to Ved. *kaví-* with its normal *i*-stem declension (*kavís*, *kavím*, etc.) has preserved the highly archaic forms nom. sg. OAv. *kauuā*, YAv. *kauua*, acc. sg. *kauuaēm* (< **kau̯ai̯am*) and nom. pl. OAv. *kāuuaiias*°.

(11) In Old Avestan there is a formal distinction of gen. du. and loc. du. (ending in *-å* < IIr. **-ās* and *-ō* < IIr. **-au̯* respectively), whereas this is unknown to Younger Avestan (with *-å* alone) as it is to Vedic (with *-os* in both cases): cfr. gen. du. OAv. *ąsaiiå* „of both shares", YAv. *vīraiiå* „of both men", *duuaiiå* „of the two", OAv. *ahuuå* „of both lifes", YAv. *bāzuuå* „of both arms", but loc. du. OAv. *zastaiiō* „in both hands", *ubōiiō* „in both", *aŋhuuō* „in both lifes".

(12) The forms of the comparative and superlative degrees of the primary (PIE.) **-i̯os-* and **-is-to-* type show various innovations in Old Indo-Aryan. Most striking is the nasal in the strong cases, OIA. *-yāṃs-*, which is missing in Avestan as in all other related languages and thus definitely is an Indo-Aryan

innovation. Moreover, there are cases in which a root-final velar is, following the form of the positive degree, not palatalized, e. g. *dīrghá-* „long" with *drā́ghīyas-*, *drā́ghiṣṭha-* in contrast to Av. *darəga-/darəya-*, but *drājiiah-*, *drājišta-*, or Ved. *raghú-* „fast" with *rághīyas-*, whereas in Avestan we find as belonging to YAv. *rayu-* „fast" regular full grade forms with palatalization, viz. *rəṇjiiah-*, *rəṇjišta-*, as one has to expect them (as the equivalents to unattested **rā́ṃhīyas-*, **rā́ṃhiṣṭha-*). A similar case of a zero-grade form instead of full-grade is Ved. *bhū́yas-* (with *bhū́yiṣṭha-*) to *bhū́ri-* „much, large, etc.", as it becomes clear from the Avestan parallel, viz. *bū̆^{i}ri-*, but *baoiiah-* (< IIr. **b^{h}áu̯H-i̯as-*). And the obviously inherited comparative and superlative degrees to PIE. **pl̥h$_1$ú-* „much" (> Ved. *purú-*, PIr. **paru-*, Gk. πολύς), YAv. *frāiiah-* „more" and *fraēšta-* „most" (cfr. Gk. πλείων, πλεῖστος), were lost in Indo-Aryan altogether except for the isolated adverb *prāyas* „for the most part" which is attested only in Epic Sanskrit[23].

(13) For the personal pronouns (a peculiarity of which is the indication of the category „number" by different stems) it is characteristic that all Indo-Aryan nominative forms have the ending *-ám* in common. By comparing the Avestan state of things it becomes clear that shorter forms without that IIr. **-ám* are underlying, which have not been continued in the Vedic language; one has to consider only, besides the longer forms advancing more and more, archaic relics like 1 du. OAv. *vā* „we two" (vs. Ved. *vā́m* < **vā-ám*), 2 sg. OAv. *tū* „you"[24] (vs. Ved. *t$_u$v-ám* = OAv. *tuuə̄m*, YAv. *tūm*, OP. *tuvam*) and 2 pl. OAv. *yūš* „you" (vs. Ved. *yūy-ám* for older **yūr-ám* = Av. *yūž-ə̄m*). Not absolutely certain is, however, OAv. *az* „I" in *as-cīt̰* „I for my part" besides *az-ə̄m* (= OP. *adam*, Ved. *ahám*).

(14) In the 1 sg. pronoun the genitive form Ved. *máma* (together with derivatives like *máma-ka-*) is opposed to OAv. *mə̄.nā* (< **mana*), YAv. *mana*, OP. *manā*, which all go back to PIr. **mana*; by the gen. form OCS. *mene* it becomes clear that this

-n-form is the older and indeed the inherited form, whereas Ved. *máma* arose only secondarily by assimilation.

(15) The oblique cases of the 1 pl. pronoun are built on PIE. **n̥s-mé* „us" (with the zero grade of **nōs/*nos* as its basis), which is reflected unaltered by PGk. **asmé* (cfr. Lesb. ἄμμε, Dor. ἀμέ) and also by PIr. **ahma*, as it is required by OAv. *ə̄hmā* and YAv. *ahma*; the form attested in the Vedic language on the other hand is acc. *asmā́n*, with the acc. pl. made plainly visible by the ending *-ān* (cfr. *áśvān* „horses")[25], in the same way as we find in Greek ἡμᾶς < Hom. ἡμέ-ας with the nominal case ending -ας being added. For the 2 pl. pronoun (Ved. *yuṣmā́n*) the development surely was similar (cfr. PGk. **usmé* [as attested in Lesb. ὔμμε, Dor. ὑμέ] < PIE. **us-mé*), but it cannot be verified, since the Avestan form is not attested.

(16) The enclitic forms in the oblique cases of the 1 pl. and 2 pl. pronouns show no difference between acc. and gen.-dat. in Ved. *nas* „us" and *vas* „you", as it is also the case in YAv. acc.-gen.-dat. *nō* < PIr. **nah* and *vō* < PIr. **u̯ah* respectively, in contrast to what we find attested in Old Avestan: 1 pl. acc. *nā̊* < PIr. **nāh* = Lat. *nōs* and 2 pl. acc. *vā̊* < PIr. **u̯āh* = Lat. *uōs* besides 1 pl. gen.-dat. *nə̄* < PIr. **nah* = Ved. *nas* and 2 pl. gen.-dat. *və̄* < PIr. **u̯ah* = Ved. *vas*. Since those lengthened grade forms demonstrably are inherited (cf. Lat. *nōs*, *uōs*, OCS. *ny*, *vy*), the Old Avestan evidence can be understood in no other way than as a relic of the original distribution of the lengthened and full grade forms in acc. and gen.-dat. respectively.

(17) For the possessive pronouns *ma-* „my", *ϑβa-* „your" and reflexive *x*^{v}*a-* „his/her own" the characteristic „pronominal" declension type is common in Old Avestan: cfr. 1, 2 sg. dat. sg. *ma-hmāi*, *ϑβa-hmāi*, 2 sg. abl. sg. *ϑβa-hmāt̰*, 1, 2 sg. loc. sg. *ma-hmī*, *ϑβa-hmī*, 2 sg. nom. pl. *ϑβōi*, and, most remarkably, nom. sg. fem. 2 sg. *ϑβōi*, refl. *x*^{v}*aē-cā* (< PIr. **ϑu̯ai̯*, **hu̯ai̯*) in agreement with Lat. *quae* etc., but in contrast to Ved. *svā́*. In Younger Avestan one small relic is preserved in loc. sg. refl. *x*^{v}*a-hmi* alone. Old Indo-Aryan, however, where as the only parallel to those

formations the reflexive possessive *svá-* „his/her own" is found, has lost any trace of the pronominal declension here.

(18) For the interrogative and indefinite pronouns PIE. **k*^{u}*e-/*k*^{u}*o-* and **k*^{u}*i-* we should expect in Indo-Iranian according to the law of the palatals **ca-*, **ka-*, and **ci-* respectively. All reflexes of palatalization are eliminated, however, in Indo-Aryan, where *k-* has been generalized apart from the isolated particle *cit* (= Av. *cīt̰*). In contrast to this a lot of *c*-forms is found in Avestan: on the one hand only OAv. gen. sg. *cahiiā* „whose" (cfr. Gk. Hom. τέο) vs. Ved. *kásya* (but also OAv., YAv. *kahiiā*, YAv. *kaŋhe*, *kahe*), YAv. dat. sg. *cahmāi* vs. Ved. *kásmai* (and OAv., YAv. *kahmāi*), loc. sg. *cahmi* vs. Ved. *kásmin* (and YAv. *kahmi*), on the other hand OAv., YAv. nom. sg. *ciš* (vs. Ved. *kís*), YAv. acc. sg. *cīm*, nom.-acc. ntr. sg. *cit̰*, pl. *cī*, and the indefinite particle *cīm* (vs. Ved. *kīm*; cfr. *mā-cim* vs. Ved. *mā́ kīm*). The same palatalization is attested also in YAv. *ca*^{i}*ti* „how many?" vs. Ved. *káti* < PIE. **k*^{u}*éti* and YAv. *cuuant-* „how long?" < PIr. **cīu̯ant-* vs. Ved. *kī́vant-*.

(19) The athematic nasal present stem PIE. **u̯i-né-d-* „to find" is preserved in unaltered form only in 3 sg. ind. act. OAv. *vīnas-tī*, YAv. *vīnas-ti*, 3 sg. opt. med. *viṇd-ī-ta*, part. med. *viṇd-āna-* and perhaps 3 sg. ind. med. *viste* < PIr. **u̯i-n-s-tai̯*. Ambiguous are forms like 3 pl. ind. act. *viṇdəṇti* and 3 pl. inj. act. *viṇdən*, however, by which as by a hinge or over a bridge the thematic forms 3 sg. inj. act. *viṇda-t̰* and the like have come into existence, as they are attested in Vedic without any exception (cfr. *vindáti*, med. *vindáte*). Those thematicized nasal infix stems, as they are found in several Indo-European languages in different variants, are obviously of secondary character.

(20) As primary endings of the 1 sg. pres. (and fut.) the Indo-European proto-language had **-mi* in athematic and **-ō* in thematic stems respectively (cfr. Gk. εἰ-μί vs. λέγ-ω). In Proto-Indo-Iranian (athematic) **-mi* has been extended then to the thematic stems, which thus are characterized by IIr. **-ā-mi*; this innovation was not carried through, however, in Old Avestan,

where original *-ā* is still found in *spasiiā* „I see" (vs. Ved. *páśyā-mi*), *pərᵊsā* „I ask" (vs. OP. *pr̥sāmiy*, Ved. *pr̥cchā́mi*) and fut. *vaxšiiā* „I shall tell" (vs. Ved. *vakṣyā-mi*) as well as in some subjunctive forms (*xšaiiā, zbaiiā*, etc.). But Old Avestan has also one example of the ending *-āmī* in *auuāmī* „I assist", from which we may conclude that this extension has developed independently in the two branches of Indo-Iranian, the Iranian one being more conservative.

(21) As regular form of the 2 sg. imperative of the verb „be" PIE. *$*h_1és$-mi* should appear *$*h_1s$-d^hí* with zero-grade of the stem (as in the corresponding forms OAv. *i-dī* „go!" = OP. *°i-diy* = Ved. *i-hí* or YAv. *jaⁱ-δi* „beat!" = OP. *ja-diy* = Ved. *ja-hí* belonging to the roots *i* „to go" and *jan* „to beat" respectively). The only evidence[26] for that expected form is OAv. *zdī* „be!". In contrast to this we find in Vedic *edhí* < **azdʰí* with the generalized full-grade, which form is an obvious innovation.

(22) The secondary ending of the 2 sg. middle PIE. **-so*, being supported by Gk. -ο (ἐφέρεο, ἔθεο, φέρ-οι-ο), **-so* (ἐγέγραψο), Lat. *-re*, is attested by Av. *-ŋhā̆/-šā̆*, OP. *-šā* (< IIr. **-sa*), but has been replaced in Old Indo-Aryan by Ved. *-thās* (perhaps < PIE. **-th_2o-es*). The evidence includes OAv. inj. pres. *paⁱrii-aoγ-žā* (< IIr. **°aug*ʰ*-sa*) „you give order", opt. pres. *xšnəuu-ī-šā* „you might satisfy", inj. aor. *mǝ̄ŋghā* (< IIr. **man-sa*) „you think"; YAv. inj. pres. *zaiia-ŋha* „you were born" (vs. Ved. *jā́ya-thās*), opt. pres. *yaz-aē-ša* = OP. *yad-ai-šā* „you might worship" (vs. Ved. *-e-thās*).

(23) That the constant long vowel in the root present and root aorist participles of roots ending in *-ā*, as shown in Ved. *pā́nt-* „drinking", *yā́nt-* „going", *bhā́nt-* „shining", fem. *bhātī́-*, etc., cannot be original, has been shown by Hoffmann (1969: 4 f. = 1975: 241 f. fn. 13). For the zero-grade we find otherwise in root aorist participles is seen also in Gk. θείς < *θέντ-ς < PIE. *$*d^h h_1$-ént-*, στάντ- < PIE. *$*sth_2$-ént-*, etc. and in Avestan. There the equivalent of Ved. *bhā́nt-*, fem. **bhātī́-* is attested, i. a., in YAv. *viiāuuantəm* < IIr. **u̯i-ā-bʰH-ant-am* and YAv. *viiāuuaⁱti* < IIr. **u̯i-ā-bʰH-at-ī*; those forms reflect clearly PIE. *$*b^h h_2$-ént-*, fem. *$*b^h h_2$-n̥t-íh_2* and are fully comparable with the Greek forms quoted above.

(24) Finally may be listed a number of Avestan lexical items, which have no Old Indo-Aryan counterpart, but must be archaisms:

(a) adj. OAv., YAv. *ā̆sna-* „nearby" (= OP. *ašna-*), going back to IIr. **azdna-*, as shown by the degrees of comparison *nazdiiah-*, *nazdišta-*, which agree with Ved. *nédīyas-*, *nédiṣṭha-*, to which for their part no positive degree exists, however;

(b) OAv. *jiiātu-* „life" < PIE. *$*g^u̯ i̯eh_3$-tu-*, which has been remodelled, however, into Ved. *jīvā́tu-* after *jīvá-* „living" (the equivalent of OP. *jīva-*, Av. *juua-* < **jīu̯a-*) and the secondary root *jīv*, most probably to avoid an association with root *jyā* „to injure, rob, deprive" (whereas the cognate verbs Av. *ziiā*, OP. *d°* for phonetic reasons did cause nothing of the sort);

(c) YAv. *taršu-* „dry" < PIE. **tr̥sú-* like Gmc. **þurzú-* „dry" is identical as to its form with Ved. *tr̥ṣú-* but not in its meaning „greedy", which came into being only by secondary development;

(d) YAv. *⁺parᵊsa-* (written *pərᵊsa-*) „piglet" (= Khot. *pā'sa-*) < PIE. **porḱo-* (cfr. Lat. *porcus*), as recognized by Hoffmann (1967: 35–37 = 1976: 491–493), whereas the Indo-Aryans have displaced this old inherited lexeme;

(e) OAv., YAv. *pasu- vīra-* „cattle and men", joined in a dual formula, which may be quoted as an archaic idiomatic expression inherited from the Indo-European proto-language (cfr. Umbr. *u(e)iro pequo*, etc.), that has been lost in Vedic, but may have given rise to *virapśá-* (< **vīra-pśv-á-*) „abundance, wealth";

(f) OAv. *dᵊbązah-*, YAv. *bązah-* „firmness, strength" < PIE. **bʰéngʰ-es-* (cfr. Gk. πάχος with irregular zero-grade), which has no equivalent OIA. **báṃhas-*, as one should expect it as belonging to *bahú-* „much, abundant" within the so-called „Caland system";

(g) YAv. *mər^{ə}ti-* „death" (= Lat. *mors*) < PIE. **mr̥-tí-*, which has been remodelled in Old Indo-Aryan into *mr̥tyú-* by crossing IIr. **mr̥tí-* with the antonym **jyā́tu-/jīvā́tu-* (see (b) above);

(h) YAv. *mər^{ə}zu°* „short" < IIr. **mr̥ǰ^{h}ú-* < PIE. **mr̥ǵ^{h}ú-* (= Gk. βραχύς) with its only cognate in Ved. adv. *múhur* which is different, however, formally (with *-u-*, notwithstanding its explanation) and semantically („suddenly, at once");

(i) YAv. *vīsa^{i}ti* „20" < PIE. **u̯ih$_{1}$ḱm̥ti-* (cfr. Gk. Dor. ϝί́κατι, Lat. *uīgintī*, etc.) in contrast with Ved. *viṃśatí-* with a secondary nasal (which should not be compared with the one in Oss. Dig. *insæj*), perhaps owing to a hypercorrect Sanskritization;

(j) YAv. *snāuuar^{ə}* „sinew" as evidence of an old heteroclitic *r/n*-stem contrary to Ved. *snā́van-* and only faint traces of the *r*-stem in Ved. *a-snāvirá-* „without sinews" and Pāli *nhāru* < **snāru-* < **snāvr°*;

(k) OAv. gen. sg. *x^{v}ə̄ṇg*, YAv. *hū* < PIr. **huu̯ánh* (from *huuar$^{ə̄}$* „sun" = Ved. *s$_{ú}$var*) and OAv. *x^{v}ə̄nuuaṇt-*, YAv. *x^{v}an°* „sunny" < PIr. **huu̯an-u̯ant-* as relics (the only ones in the Indo-Iranian languages at all) of *n*-forms of the old heteroclitic paradigm.

Not much store should be set on phenomena lost in Old Indo-Aryan, but retained in Iranian such as, e. g., the nouns OAv., YAv. *spəṇta-* „beneficient, prosperous, holy" (< IIr. **ću̯anta-*; cfr. OCS. *svętъ* etc.), YAv. *mar^{ə}za-* „march, border zone" and the verbal roots YAv. *vad* „to lead", *part* „to fight" or *varz* „to work"; here should be mentioned also the suffix IIr. **-ai̯na-* used for adjectives indicating material composition (like OP. *aϑang-ai̯na-* „of stone", YAv. *aiiaŋh-aēna-* „of metal", *ər^{ə}zat-aēna-* „of silver", etc.), which in Old Indo-Aryan has got lost.

The topic of those Iranian archaisms that contrast sharply with Vedic innovations only rarely has been discussed in the relevant literature, only Mayrhofer (1997: 119 f.) being an exception. When we look after the conclusion which we must draw from this situation for the reconstruction of the Indo-Iranian

proto-language and for the comparative study of the Indo-European languages in general, it will become absolutely clear that for proving the Indo-Iranian origin of some linguistic phenomenon it is neither sufficient nor adequate to refer to the Indo-Aryan (Vedic and Sanskrit) data alone, notwithstanding the abundant, very old and usually rather faithful textual tradition of this language, to which the fragmentary Old Iranian records in the form of two limited corpora of quite monotonous and stereotyped texts are diametrically opposed. For both Avestan and Old Persian present for many phonological and most morphological and other phenomena so few attestations that it frequently is quite difficult to decide whether we have to do with an inherited form, with an archaism or an innovation. The problems with the often rather ambiguous Old Persian cuneiform writing on the one hand and the cleverly devised Avestan script on the other hand, which renders all phonetic subtleties, but thereby is more obscuring than inspiring, increase those difficulties. Thus the Old Iranian texts at first glance seem to be put in the shade of the Vedic texts.

In reality, however, the situation is not so hopeless, as one would think. One only has to proceed cautiously step by step and to scrutinize every form philologically and linguistically with utmost thoroughness, and in any case, in order to do not rush things, the first task must be to reconstruct by referring equally to both the Indo-Aryan and the Iranian evidence the situation at the common Indo-Iranian linguistic stage, since only this is on the same level as the other members of the Indo-European language family. In plain words this means, that the Indo-Aryan evidence, and even the Rigvedic one, must first be examined for its originality by comparing it with the Iranian data alone in view of their equal rights. Apart from the wealth and diversity of the textual evidence written in younger Iranian languages, the archaism of the Old Avestan language makes it only too clear, as we have seen, that the Old Indo-Aryan language despite the old age of the earliest Vedic texts is not the only solid fundament

within the Indo-Iranian area, by which we can support the reconstruction of Proto-Indo-European.

In the same way as one can conclude from Middle Indo-Aryan archaisms, which are in contrast to Vedic forms, that these are dialectal traits of unremembered Old Indo-Aryan dialects, which have gone away, we may infer also that Iranian (esp. Old Avestan) archaisms are evidence for dialectal diversity of Proto-Indo-Iranian. So at best we may conclude that this was not a homogeneous language spoken everywhere in the same form, but rather a continuum of closely related or almost identical dialects. In any case, however, one must not ignore those linguistic features that differ from the Vedic and Sanskrit evidence and after all must come to the result that Old Indo-Aryan is not identical with Proto-Indo-Iranian and is not so close to Proto-Indo-European as many people maintain. This remark is aimed in particular at the adherents (mostly ignorant in historical comparative linguistics) of the 'Out-of-India Theory', that in the last few decades came back with a vengeance and has been dealt with carefully by Bryant 2001 and by some[27] of the contributors to Bryant and Patton 2005.

Endnotes

---

[1] The following abbreviations may be noted: Av. = Avestan, Gk. = Greek, Gmc. = Germanic, IIr. = Indo-Iranian, Lat. = Latin, MP. = Middle Persian, NHG. = New High German, OAv. = Old Avestan, OCS. = Old Church Slavonic, OIA. = Old Indo-Aryan, OIr. = Old Iranian, OLat. = Old Latin, OP. = Old Persian, Oss. = Ossetic, PGk. = Proto-Greek, PIE. = Proto-Indo-European, PIr. = Proto-Iranian, Ved. = Vedic, YAv. = Younger Avestan.

[2] From this source are descended the Nuristani languages, too, which are spoken on both sides of the frontier between Afghanistan and Pakistan. Because they have been recorded only from the beginning of the 20th century, their classification (as a third branch of Indo-Iranian or as an archaic offshoot of either Indo-Aryan or Iranian) is still a matter of dispute; see most recently the discussion in Degener 2002.

[3] The fundamental principle maintained by Schlerath (1981: 186 = 2000: 334), that in a reconstructed language dialects cannot exist („in einem Rekonstrukt kann es keine Dialekte geben"), seems quite naive to me, as if reconstructed languages were closer to the ideal of a uniform standard type than natural languages, particularly since by reconstruction we do not always reach the same chronological level and the same spatial distribution.

[4] For the problems concerning such intermediate units, as they have been postulated on various occasions, see recently Hock 2000.

[5] See YAv. *kaoiiąm* < **kau̯i̯-ām* (vs. Ved. *kavīnā́m* „seers"); *yāϑβąm* < **yāϑu̯-ām* (vs. *yātunąm* „sorcerers") etc., which relics may be registered also as archaisms preserved in the Younger Avestan language.

[6] Cfr., e. g., Meillet (1908: 25–27); Schmitt (1989: 2; 2000: 8 f.) and most recently Tremblay (2005b: 674 fn. 3).

[7] Of little or no relevance for Proto-Indo-Iranian is the (at least partial) coincidence of PIE. **l* and **r*, notwithstanding the evidence available for /l/ in both Iranian and Indo-Aryan languages: cfr. Mayrhofer 2002.

[8] Evidence for this „frequentative" is found in Avestan (e. g., *aδaxšaiiaētā* „he used to instruct" or, without augment, *nišhiδōiš* „you used to sit down"), Old Persian (*akunavayantā* „they used to bring about"), Middle Persian, Sogdian, Khotanese, and Choresmian. Likewise a feature typical of Iranian is the use of the perfect optative for expressing the irrealis of the past (cfr. OP. *naiy ... kašciy, haya ... caxriyā* „not anyone, who might have done").

[9] A systematic description of Proto-Indo-Iranian is missing up to this day; at hand there are only more or less systematic accounts comparing the historically attested ancient languages, esp. that by Bartholomae 1895–1901, and, most recently, the concise sketch by Skjærvø (2006: 371–376).

[10] This holds true also when taking into account the Middle Indo-Aryan archaisms – like Pāli *idha* „here" instead of Ved. *ihá* (= OAv., OP. *idā*, YAv. *iδa*) and many others – which are in contrast to Vedic forms with a more advanced phonological development and by this indicate that there existed

„non-Sanskritic" dialects in Old Indo-Aryan times: cfr., e. g., Mayrhofer 1955; Oberlies 1999.

[11] This does not mean that there are no Indo-Aryan innovations in morphology and syntax. Here I mention only the case ending of the nom.-acc. pl. of neuter *a*-stems Ved. *-āni* (like *bhúvanāni* „beings") or the indeclinable so-called gerunds in *-tvā* and *-ya*.

[12] The influences of the non-Indo-Aryan on the Indo-Aryan languages were often discussed; a quite significant study of that topic we have in Kuiper 1967, by which the phenomenon of the Indic linguistic area (*Sprachbund*) mirrored in those influences for the first time received a more historical dimension.

[13] A phonological development like that of *mīḍhá-* is of particular value as evidence, because the process itself on no account is reversible. A still more illustrative example may be found in Ved. *vóḍhar-* „draught animal", which is an agent noun in PIE. **-tér-/-tor-* to the root PIE. **u̯eǵʰ* (> Ved. *vah* „to drive, draw" etc.), i. e. PIE. **u̯éǵʰ-tor-* (cfr. Lat. *uector* „draught animal"). Whereas on the one hand Ved. *vóḍhar-* developed owing to Bartholomae's law via IIr. **u̯ájdʰar-* and pre-Ved. **vážd ʰar-*, in the Avestan language a synonymous *vaštar-* is found instead of **važdar-* by analogical levelling. If one would start from the Vedic form *vóḍhar-*, however, as the 'Indigenous (Indo-)Aryanists' might do it, no one could explain the Avestan form (let alone Lat. *uector*), since a phonological change of **o* (or **au̯*) to *aš* (or sim.) is entirely unimaginable.

[14] A difference like that in the numeral „one" – the isolated OIA. *éka-* (and already in Near Eastern Indo-Aryan *a-i-ka°* = **aika-*) vs. Iran. **ai̯u̯a-* (OAv., YAv. *aēuua-*, OP. *aiva-*, etc.), which is matching Gk. οἶϝος „alone" – may not simply be explained as archaism and innovation respectively, but probably reflects a Proto-Indo-Iranian dialectal divergence.

[15] I did not strive for completeness and gave more attention to correspondences (i. e. morphologically identical word-forms) than to similarities (i. e. similar, not exactly corresponding word-forms derived from the same root).

[16] There is no lack of archaisms, however, found only in Middle Iranian (but neither in Old Iranian nor in Vedic), esp. in Khotanese; cfr. Tremblay (2005a: 3b). But there are Younger Avestan archaisms, too, in contrast to Old Avestan innovations, for which cases one should not overlook the fact that Old Avestan was still living when the texts were composed, whereas Younger Avestan was used by Zoroastrians already rather early as a merely traditional language.

[17] The decisive structural difference between the various Old Iranian languages is that Younger Avestan and Old Persian in practice have replaced the inherited aspectual system of the verb by a tense system. Moreover, in the ordinal numbers YAv. *ϑritiia-*, OP. *çitīya-* „third" (< PIr. **ϑritīya-*) archaic IIr. **tr̥tī́ya-* (as required by Ved. *tr̥tī́ya-*, Pāli *tatīya-*, etc.) is replaced by **tri°* > **ϑri°* in analogy to the stem of the cardinal PIE., IIr. **tri-* „three". And it may be added, too, that in Younger Avestan the abl. sg. has the case ending *-at̰* throughout,

whereas in Old Avestan the abl. sg. is mostly identical with the gen. sg. and has a particular ending, *-āt̰* < **°a-at̰*, only in the *a*-stems.

[18] In *tā* and in YAv. *tūiriia-* „father's brother" = Ved. *pitr̥vyà-* we see apparently an optional simplification of the initial cluster *pt-*.

[19] The readings of several „good" manuscripts are in favour of the view that the form of acc. sg. and nom. pl. represent *pt-*: cfr. de Vaan (2003: 538 with fn. 702).

[20] See the preceding note.

[21] Regardless of this we find in the same paradigm also secondary forms (apparently analogically built upon the nom. sg.) in instr. pl. YAv. *vaγžibiš* (vs. Ved. *vāg-bhís*), abl. pl. YAv. *vaγžibiiō*, instr. du. YAv. *vaγžibiiā°*.

[22] In Old Persian the infinitive *cartanaiy* also has *c-* and by this is evidence of the underlying full-grade of the verbal root in this infinitive formation.

[23] One may compare the isolated superlative degree OP. (adv.) *duvaištam* „very far off; for a very long time to come" (< PIE. **du̯eh₂-isto-*), that is built on the full grade like Gk. δήν (< *δϝᾱν), δηρός etc., whereas Ved. *dávīyas-*, *daviṣṭhám* follow the quite regular pattern of *śū́ra-*, *śáviṣṭha-*.

[24] The pronominal character of *tū* has been doubted by Strunk (1975: 320–324 = 2005: 347–351), who saw there also the particle Av. *tū* „now, but".

[25] In all probability this remodelling did not take place on the direct way; pre-Ved. **asmá* on the contrary may first have been enlarged – cfr. (13) above – by the passepartout **-ám* to **asmā́m* and only this form by dissimilation may have been brought into line with the nominal ending of the acc. case.

[26] Problematic (and rather enigmatic indeed) is the equivalent found in Greek, ἴσ-θι.

[27] This does not apply to Misra 2005, where, apart from other idiosyncratic views, the Old Iranian languages in all seriousness and in the absence of a sound methodology are put on the same level as Middle Indo-Aryan.

References

Bader, Françoise (ed.)

1997 *Langues indo-européennes*, 2nd edn. Paris: CNRS Editions.

Bartholomae, Chr.

1895–1901 Vorgeschichte der iranischen Sprachen, in: Geiger and Kuhn (1895–1901), 1–151.

Benveniste, E., and Renou, L.

1934 *Vṛtra et Vṛθragna. Étude de mythologie indo-iranienne.* Paris: Imprimerie Nationale.

Bogoljubov, M. N. (ed.)

2006 *Indoiranskoe jazykoznanie i tipologija jazykovyh situacij. Sbornik statej k 75-letiju professora A. L. Grjunberga.* Sankt-Peterburg: Nauka.

Bryant, Edwin

2001 *The Quest for the Origins of Vedic Culture. The Indo-Aryan Migration Debate.* Oxford: Oxford University Press.

Bryant, Edwin F. and Patton, Laurie L. (eds.)

2005 *The Indo-Aryan Controversy. Evidence and inference in Indian history.* London and New York: Routledge.

Darms, Georges

1978 *Schwäher und Schwager, Hahn und Huhn. Die Vṛddhi-Ableitung im Germanischen.* München: Kitzinger.

Degener, Almuth

2002 The Nuristani Languages, in: Sims-Williams (2002), 103–117.

de Vaan, Michiel

2003 *The Avestan Vowels.* Amsterdam and New York: Rodopi.

Geiger, Wilh., and Kuhn, Ernst (eds.)

1895–1901 *Grundriss der Iranischen Philologie. Erster Band. 1. Abteilung.* Strassburg: Trübner.

Hock, Wolfgang

2000 Balto-Slavisch, Indo-Iranisch, Italo-Keltisch. Kriterien für die Annahme von Sprachgemeinschaften in der Indogermania, in: Range (2000), 119–145.

Hoffmann, Karl

1967 Drei indogermanische Tiernamen in einem Avesta-Fragment, *Münchener Studien zur Sprachwissenschaft* 22: 29–38; reprinted in: Hoffmann (1976), 486–493.

1969 Vedisch *vidh, vindh*, *Die Sprache* 15: 1–7; reprinted in: Hoffmann (1975), 238–244.

1975 *Aufsätze zur Indoiranistik. Band 1.* Wiesbaden: Reichert.

1976 *Aufsätze zur Indoiranistik. Band 2.* Wiesbaden: Reichert.

Kellens, Jean

2006 Les „Observations critiques“ de Meillet aujourd'hui, in: Bogoljubov (2006), 276–281.

Kuiper, F. B. J.

1942 *Notes on Vedic Noun-inflexion.* Amsterdam: Noord Hollandsche Uitgevers.

1967 The Genesis of a Linguistic Area, *Indo-Iranian Journal* 10: 81–102; reprinted in: Kuiper (1997), 78–99.

1997 *Selected Writings on Indian Linguistics and Philology.* Amsterdam and Atlanta, GA: Rodopi.

Mayrhofer, Manfred

1955 Das Problem indogermanischer Altertümlichkeiten im Mittelindischen, in: Spies (1955), 219–241.

1996 *Etymologisches Wörterbuch des Altindoarischen. II. Band.* Heidelberg: Winter.

1997 L'indo-iranien, in: Bader (1997), 101–122.

2002 Zur Vertretung der indogermanischen Liquiden in den indo-iranischen Sprachen, *Indologica Taurinensia* 28: 149–161.

2005 *Die Fortsetzung der indogermanischen Laryngale im Indo-Iranischen.* Wien: Österreichische Akademie der Wissenschaften.

2006 *Einiges zu den Skythen, ihrer Sprache, ihrem Nachleben.* Wien: Österreichische Akademie der Wissenschaften.

Meillet, A.

1908 *Les dialectes indo-européens.* Paris: Champion.

1917 Observations critiques sur le texte de l'Avesta, *Journal Asiatique* 1917, II: 183–214.

Meiser, Gerhard, and Hackstein, Olav (eds.)

2005 *Sprachkontakt und Sprachwandel. Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.–23. September 2000, Halle an der Saale*. Wiesbaden: Reichert.

Misra, Satya Swarup

2005 The Date of the Rigveda and the Aryan Migration. Fresh linguistic evidence, in: Bryant and Patton (2005), 181–233.

Monna, Maria Cornelia

1978 *The Gathas of Zarathustra. A Reconstruction of the Text* (Proefschrift Rijksuniversiteit te Leiden). Amsterdam: Rodopi.

Oberlies, Thomas

1999 Middle Indo-Aryan and (the) Vedic (Dialects) (Miscellanea Palica VII), *Historische Sprachforschung* 112: 39–57.

Range, Jochen D. (ed.)

2000 *Aspekte baltistischer Forschung*. Essen: Die Blaue Eule.

Rix, Helmut (ed.)

1975 *Flexion und Wortbildung. Akten der V. Fachtagung der Indogermanischen Gesellschaft, Regensburg, 9.–14. September 1973*. Wiesbaden: Reichert.

Schlerath, Bernfried

1981 Ist ein Raum/Zeit-Modell für eine rekonstruierte Sprache möglich?, *Zeitschrift für Vergleichende Sprachforschung* 95: 175–202; reprinted in: Schlerath (2000), 323–348.

2000 *Kleine Schriften. Band I*. Dettelbach: Röll.

Schmitt, Rüdiger (ed.)

1989a *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert.

Schmitt, Rüdiger

1989b Iranische Sprachen: Begriff und Name, in: Schmitt (1989a), 1–3.

2000 *Die iranischen Sprachen in Geschichte und Gegenwart*. Wiesbaden: Reichert.

2003 *Meno-logium Bagistano-Persepolitanum. Studien zu den altpersischen Monatsnamen und ihren elamischen Wiedergaben*. Wien: Österreichische Akademie der Wissenschaften.

Sims-Williams, Nicholas (ed.)

2002 *Indo-Iranian Languages and Peoples* (Proceedings of the British Academy, 116). Oxford: Oxford University Press.

Skjærvø, Prods Oktor

2006 Iran. vi. Iranian Languages and Scripts, in: Yarshater (2006), 344–377.

Spies, Otto (ed.)

1955 *Studia Indologica. Festschrift für Willibald Kirfel*. Bonn: Orientalisches Seminar.

Strunk, Klaus

1975 Horizontale Verkettung und vertikale Flexion in indogermanischen Sprachen, in: Rix (1975), 296–324; reprinted in: Strunk (2005), 323–351.

2005 *Kleine Schriften. 1. Teilband*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Tichy, Eva

1985 Avestisch *pitar-/ptar*. Zur Vertretung interkonsonantischer Laryngale im Indoiranischen, *Münchener Studien zur Sprachwissenschaft* 45: 229–244.

Tremblay, Xavier

2005a Iranian Historical Linguistics in the Twentieth Century, *Indo-European Studies Bulletin* 11, 1: 1–23.

2005b Bildeten die iranischen Sprachen ursprünglich eine genetische Familie oder einen Sprachbund innerhalb des indo-iranischen Zweiges? Beiträge zur vergleichenden Grammatik der iranischen Sprachen V, in: Meiser and Hackstein (2005), 673–688.

Yarshater, Ehsan (ed.)

2006 *Encyclopædia Iranica. Volume XIII*. New York, NY: Encyclopædia Iranica Foundation.